TP 14569E

Work Zone Traffic Management on Highways

Prepared for Transportation Development Centre
and
ITS Office of Transport Canada

by
Synectics Transportation Consultants Inc.
University of Toronto

February 2006

106-07255/

TP 14569E

Work Zone Traffic Management on Highways

by
C. Jacob, A. Hadayeghi, B. Abdulhai and B. Malone
Synectics Transportation Consultants Inc.
University of Toronto

February 2006

This report reflects the views of the authors and not necessarily the official views or policies of Transport Canada or the co-sponsoring organizations.

Transport Canada and the co-sponsoring agencies do not endorse products or manufactures. Trade or manufacturers' names appear in this report only because they are essential to its objectives.

Since some of the accepted measures in the industry are imperial, metric measures in the industry are imperial, metric measures are not always used in this report.

Un sommaire français se trouve avant la table des matières.

© 2006 Transport Canada

Transport Canada / Transports Canada

PUBLICATION DATA FORM

1. Transport Canada Publication No.	2. Project No.	3. Recipient's Catalogue No.
TP 14569E		

4. Title and Subtitle	5. Publication Date
Work Zone Traffic Management on Highways	February 2006
	6. Performing Organization Document No.

7. Author(s)	8. Transport Canada File No.
C. Jacob, A. Hadayeghi, B. Abdulhai, and B. Malone	2450-GP018

9. Performing Organization Name and Address	10. PWGSC File No.
Synectics Transportation Consultants Inc. 36 Hiscott Street, Suite 200 St. Catharines, Ontario L2R 1C8	052ss.T8663-030015
	11. PWGSC or Transport Canada Contract No. T8663-030015/002/SS

12. Sponsoring Agency Name and Address	13. Type of Publication and Period Covered
ITS Office – Transport Canada Place de Ville, Tower C, Floor 27 330 Sparks Street Ottawa, Ontario Canada K1A 0N5	Final
	14. Project Officer Pierre Bolduc

15. Supplementary Notes (Funding programs, titles of related publications, etc.)

This project is part of Canada's Intelligent Transportation Systems (ITS) R&D Plan, *Innovation Through Partnership*, funded by the ITS Office of Transport Canada under the Strategic Highway Infrastructure Program (SHIP). The Transportation Development Centre of Transport Canada served as technical authority for this project.

16. Abstract

The focus of this study was on the development of functional and system requirements for an integrated, real-time temporary condition **T**raffic **M**anagement **S**ystem for work zones on **Can**adian roadways (**TMS-Can**). TMS-Can uses a combination of electronic sensors, software, wireless and cable communications networks, and electronic signs and variable message signs to manipulate traffic flow. The travel information on traffic routing displayed by these variable message signs is computed by a computer agent driven by Reinforcement Learning. Reinforcement Learning is an approach whereby the control agent directly learns to map sensed system states to optimal actions. A simple but powerful reinforcement learning method known as Q-learning is used. A micro simulation tool – Paramics – was utilized to train the agent in an offline mode within a simulation environment in order to make it ready for field implementation.

The approach developed in this research was rigorously evaluated under simulated conditions. Results from the simulation are very encouraging and have demonstrated the effectiveness and superiority of the technique in reducing congestion in work zones. Deployment of the developed system at the field trial stage was not completed under this study.

17. Key Words	18. Distribution Statement
Work zone traffic management, TMS-Can, automatic traffic control, adaptive traffic control, intelligent transportation systems, artificial intelligence, machine learning, reinforcement learning, Q-learning	Limited number of print copies available from the Transportation Development Centre. Also available online at www.tc.gc.ca/tdc/menu.htm

19. Security Classification (of this publication)	20. Security Classification (of this page)	21. Declassification (date)	22. No. of Pages	23. Price
Unclassified	Unclassified	—	xvi, 62, app	Shipping/ Handling

Transports Canada / Transport Canada

FORMULE DE DONNÉES POUR PUBLICATION

1. N° de la publication de Transports Canada	2. N° de l'étude	3. N° de catalogue du destinataire
TP 14569E		

4. Titre et sous-titre	5. Date de la publication
Work Zone Traffic Management on Highways	Février 2006
	6. N° de document de l'organisme exécutant

7. Auteur(s)	8. N° de dossier - Transports Canada
C. Jacob, A. Hadayeghi, B. Abdulhai et B. Malone	2450-GP018

9. Nom et adresse de l'organisme exécutant	10. N° de dossier - TPSGC
Synectics Transportation Consultants Inc. 36 Hiscott Street, Suite 200 St. Catharines, Ontario L2R 1C8	052ss.T8663-030015
	11. N° de contrat - TPSGC ou Transports Canada T8663-030015/002/SS

12. Nom et adresse de l'organisme parrain	13. Genre de publication et période visée
Bureau STI – Transports Canada Place de Ville, Tour C, 27e étage 330, rue Sparks Ottawa, Ontario Canada K1A 0N5	Final
	14. Agent de projet Pierre Bolduc

15. Remarques additionnelles (programmes de financement, titres de publications connexes, etc.)

Ce projet fait partie du Plan de R&D du Canada sur les systèmes de transports intelligents (STI), *Innover par l'établissement de partenariats*, financé par le Bureau des STI ainsi que par Transports Canada dans le cadre du programme stratégique d'infrastructures routières (PSIR). Le Centre de développement des transports de Transports Canada a agi comme responsable technique pour ce projet.

16. Résumé

La présente étude a porté sur l'élaboration des exigences fonctionnelles et la définition des caractéristiques d'un système intégré de gestion de la circulation en temps réel dans les zones de travaux routiers (***Traffic Management System for work zones on Canadian roadways*** : **TMS Can**). Le TMS-Can utilise une panoplie de capteurs électroniques, de logiciels, de réseaux de communications sans fil et par câble, de tableaux d'affichage électroniques et de panneaux à message variable pour gérer la circulation. Un agent logiciel apprenant par renforcement calcule l'itinéraire affiché sur les panneaux à message variable. Dans l'apprentissage par renforcement, l'agent apprend à traiter des perceptions pour agir de façon optimale sur l'environnement. Le système utilise un type d'apprentissage par renforcement simple mais puissant, appelé apprentissage Q. Un outil de microsimulation – Paramics – a été utilisé pour la formation hors ligne de l'agent dans un environnement simulé, en préparation à sa mise en œuvre sur le terrain.

La méthode élaborée au cours de la présente recherche a été rigoureusement évaluée dans des conditions simulées. Les résultats de cette simulation sont très encourageants. Ils démontrent en effet que la technique est capable mieux que toute autre de réduire la congestion dans les zones de travaux. Le système développé n'a pas été mis à l'essai en situation réelle dans le cadre de cette étude.

17. Mots clés	18. Diffusion
Gestion de la circulation dans les zones de travaux, TMS-Can, régulation automatique de la circulation, régulation adaptative de la circulation, systèmes de transports intelligents, intelligence artificielle, apprentissage machine, apprentissage par renforcement, apprentissage-Q	Le Centre de développement des transports dispose d'un nombre limité d'exemplaires imprimés. Disponible également en ligne à www.tc.gc.ca/cdt/menu.htm

19. Classification de sécurité (de cette publication)	20. Classification de sécurité (de cette page)	21. Déclassification (date)	22. Nombre de pages	23. Prix
Non classifiée	Non classifiée	—	xvi, 62, ann.	Port et manutention

CDT/TDC 79-005
Rev. 96

Canada

ACKNOWLEDGEMENTS

The cooperation of the Ontario Ministry of Transportation, especially Mr. Phil Masters and his staff, as well as the support and advice of M. Pierre Bolduc, Transportation Development Centre of Transport Canada, are gratefully acknowledged.

We are grateful to the Project Steering Committee for its valuable comments and insights throughout the research.

Steering Committee:

Phil Masters, Ministry of Transportation
Rod Edwards, Harpar Management
Bruce Zvaniga, City of Toronto
Edward Soldo, Region of Halton
Madeleine Betts, Transport Canada

Project Team

Brian Malone, Co-Project Manager
Baher Abdulhai, Co-Project Manager
Alireza Hadayeghi, Project Engineer, Methodology
Celine Jacob, Project Engineer, Model Development

EXECUTIVE SUMMARY

Canada's road network is maturing. The increasing growth in traffic and limitations on the expansion of existing infrastructure is taking a toll on the existing transportation network, accelerating the ageing process. Therefore, there is an increasing need for maintenance and rehabilitation of our roadways. This translates into more work zones on highways accompanied by more traffic congestion.

Traditional work zone traffic control focuses on advance warning, site identification and driver guidance. The main purpose of these efforts is to protect workers from traffic, and to allow road users to navigate the work zone safely and efficiency. Over the past several years, many techniques and/or programs have been implemented to enhance work zone safety and facilitate traffic control. These include innovations such as work zone intrusion alarms, portable rumble strips, flashing stop/slow paddles, barrier lighting units, etc. While these measures are an enhancement over traditional traffic control measures, their inability to dynamically direct and/or respond to changing traffic conditions results in lost opportunities to increase mobility and economic productivity, as well as to reduce costs and environmental impacts. Therefore, there exists a need to provide reliable real-time routing information to drivers as they are about to enter the work zone.

The main focus of this project was on the application of Intelligent Transportation Systems to work zone traffic management on highways, for the purposes of promoting safety, relieving congestion, improving mobility, and minimizing negative impacts due to transportation. This research introduced the use of Reinforcement Learning, an Artificial Intelligence method from Machine Learning, to provide optimal control for a freeway-arterial or Express/collector corridor with work zones. It presented the methodology, development, and simulated testing and results of the Artificial Intelligence based computer agent, in preparation for real time implementation in the field. The approach focused on providing real time route recommendations that could be relayed to motorists through Variable Message Signing (VMS) in order to direct traffic effectively to minimize system-wide delay and congestion due to construction.

Reinforcement Learning is an approach whereby a computer control agent directly learns optimal strategies via feedback reward signals from its environment. A simple but powerful Reinforcement Learning method known as Q-Learning is used. A micro simulation tool – Paramics has been used to train the agent in an offline mode in a simulated environment to make it ready for field implementation.

The developed system, named as TMS-Can, uses a control strategy driven by the Reinforcement Learning Agent to provide traffic management capabilities and traveler information on traffic routing, and detours to drivers. With real-time information, drivers can be aided in route selection, allowing them to optimize the available network capacity restrictions imposed due to work zone construction. Simulated experiments with the Learning Agent conducted for the study area in the peak morning period with existing traffic conditions have shown considerable improvement over the base case of current

traffic status. The results of simulation tests have proved that a real-time dynamic like the Q-Learning approach can provide much better control than that of fixed diversion.

After developing the software component of the TMS-Can system and obtaining promising results from simulated conditions, the next step of this assignment was to be implementation of the system in real-time traffic in the actual work zone at the trial site. The objective of the field trial was to achieve proof of the concept in a real world situation. Deployment at the field trial stage, however, was not completed due to unexpected completion of construction in the work zone that had been used as the basis for the simulations. Attempts to identify and deploy the TMS-Can at an alternate site were unsuccessful.

Lack of field deployment made it impossible to verify the potential improvements in reducing traffic delay that were identified as being possible. Unsuccessful attempts to select an alternate test location highlighted a drawback in the existing approach, that optimal site conditions were needed to achieve setup and deployment. While an optimum site was identified for the initial trials, redeployment was not easily achieved, as alternates with similar conditions were not identified. This fact was complicated by the requirement to coordinate physical site conditions, including parallel alternate routes and vehicle detection capability, with ongoing construction activity at a suitable site.

Despite the failure to achieve a field trial of the system developed, simulated tests were promising and subsequent field deployment could be undertaken, achieving proof of concept.

SOMMAIRE

Le réseau routier du Canada vieillit. La croissance du trafic et les limites à l'expansion de l'infrastructure existante accélèrent le vieillissement du réseau. Les routes ont donc de plus en plus besoin d'être entretenues et remises à neuf. Il s'ensuit une multiplication des zones de travaux, et, corollairement, des congestions routières.

La gestion de la circulation dans les zones de travaux consiste ordinairement à prévenir les usagers des travaux en cours, à identifier la zone de travaux et à guider les conducteurs. Ces mesures visent principalement à protéger les travailleurs du danger de la circulation, et à aider les usagers de la route à franchir la zone des travaux en toute sécurité. Ces dernières années, de nombreuses techniques et/ou programmes ont été mis en œuvre afin d'accroître la sécurité dans les zones de travaux et de faciliter la gestion de la circulation. Des innovations sont notamment apparues, comme les alarmes d'intrusion dans les zones de travaux, les bandes rugueuses portables, les panneaux arrêt/ralentir clignotants, les modules d'éclairage de barrières, etc. Ces mesures représentent certes une amélioration par rapport aux mesures classiques de gestion de la circulation. Mais leur incapacité à diriger de façon dynamique la circulation et/ou à réagir en temps réel au changement des conditions de circulation changeantes fait en sorte que l'on rate des occasions de maximiser la mobilité et la productivité, et de réduire les coûts et les impacts environnementaux. Il existe donc un besoin de transmettre aux conducteurs qui s'approchent de la zone des travaux une information fiable et en temps réel concernant les itinéraires alternatifs.

Le présent projet portait principalement sur l'application de systèmes de transports intelligents à la gestion de la circulation dans les zones de travaux routiers, aux fins de promouvoir la sécurité, de soulager la congestion, d'améliorer la mobilité et de minimiser les effets négatifs des transports sur l'environnement. Cette recherche a innové en ce qu'elle a fait appel à l'apprentissage par renforcement, une branche de l'intelligence artificielle associée à l'apprentissage machine, pour optimaliser la gestion de la circulation dans un corridor autoroutier/artériel ou collecteur/express qui comprend des zones de travaux. Elle a présenté la méthodologie, le développement et l'essai par simulation de l'agent logiciel à intelligence artificielle, ainsi que les résultats de cet essai préparatoire à une mise en œuvre en situation réelle. L'approche consistait à calculer en temps réel des itinéraires de rechange et à relayer cette information aux automobilistes par des panneaux à message variable (PMV), afin de diriger efficacement la circulation et de minimiser les retards et les congestions dus aux travaux dans l'ensemble du réseau routier.

L'apprentissage par renforcement est une approche par laquelle un agent logiciel apprend des stratégies optimales en captant les signaux de renforcement transmis par son environnement. Le système utilise un type d'apprentissage par renforcement simple mais puissant, appelé apprentissage-Q. Un outil de microsimulation – Paramics – a servi à former l'agent en mode hors ligne dans un environnement simulé pour le préparer à sa mise en œuvre sur le terrain.

Le système développé, baptisé TMS-Can, utilise une stratégie définie par l'agent apprenant par renforcement pour gérer la circulation et transmettre aux conducteurs l'information sur les itinéraires alternatifs. Comme l'information leur est communiquée en temps réel, les conducteurs sont mieux à même de choisir le meilleur itinéraire, ce qui permet d'optimiser la capacité du réseau dans les limites des restrictions imposées par les zones de travaux. Des simulations de la zone d'étude dans des conditions de circulation typiques de l'heure de pointe du matin ont révélé que le recours à l'agent apprenant améliorait considérablement la situation par rapport à l'état actuel de la circulation. Les résultats de ces simulations ont montré qu'une approche dynamique comme celle qui s'appuie sur l'apprentissage-Q peut donner des résultats de beaucoup supérieurs à ceux obtenus avec les dispositifs fixes de déviation de la circulation.

Après avoir développé le logiciel du système TMS-Can et obtenu des résultats prometteurs dans des conditions simulées, l'équipe de recherche s'est préparée à mettre en œuvre le système en situation réelle, dans la zone de travaux étudiée. L'objectif de cet essai en situation réelle était de valider le concept dans des conditions représentant la «vraie vie». Mais ce déploiement n'a pu avoir lieu, en raison de l'achèvement hâtif des travaux dans la zone qui avait servi aux simulations. On a tenté sans succès de trouver une autre zone propice au déploiement du TMS-Can.

Faute d'un déploiement en situation réelle, il a été impossible de confirmer dans les faits que le système était supérieur aux dispositifs actuels pour réduire les congestions. Par ailleurs, les tentatives infructueuses de trouver un autre site pour l'essai ont mis en lumière une faille dans l'approche existante, à savoir que l'on avait besoin de conditions optimales pour installer et déployer le système. Un site optimal a bien été trouvé pour les premiers essais, mais on n'a pu passer à la seconde phase de déploiement, faute de trouver d'autres sites présentant des conditions similaires. La situation était d'autant plus compliquée que pour convenir à l'essai, le site devait présenter certaines conditions physiques, notamment des itinéraires alternatifs parallèles et un système de détection des véhicules, et en plus, des travaux devaient y être en cours.

Malgré l'incapacité de procéder à un essai en situation réelle, les essais simulés se sont avérés prometteurs et il est toujours possible d'envisager un déploiement sur le terrain pour valider le concept.

CONTENTS

1 INTRODUCTION 1
- 1.1 Background 1
- 1.2 Research Objective 2
- 1.3 Organization of Report 2

2 LITERATURE REVIEW 4
- 2.1 Specific Systems 4
  - 2.1.1 Travel Time Prediction System for Freeway Work Zones 4
  - 2.1.2 Real-time Advanced Warning and Traffic 6
  - 2.1.3 Portable Traffic Management System 7
  - 2.1.4 Automated Work Zone Information Systems 10
  - 2.1.5 Condition-Responsive Work Zone Traffic Control System 13
  - 2.1.6 Intelligent Transportation Systems in New Mexico 13
- 2.2 Other ITS Work Zone Systems 14
  - 2.2.1 Dynamic Work Zone Safety System 14
  - 2.2.2 SpeedGuard System 14
  - 2.2.3 Wizard CB Alert System 15
  - 2.2.4 Work Zone Intrusion Alarms 15

3 SYSTEM REQUIREMENTS FOR TMS-Can 16
- 3.1 Software Component 16
- 3.2 Hardware Component 16
- 3.3 TMS-Can Architecture 17

4 SYSTEM FUNCTIONALITY 18
- 4.1 Study Area 18
- 4.2 Field Visit 19
- 4.3 Software Development 19
- 4.4 The Reinforcement Learning Based Approach 22
  - 4.4.1 Paramics Simulator 22
  - 4.4.2 Reinforcement Learning 27

5 TRAVEL DEMAND DATA 38

6 IMPLEMENTATION OF DEVELOPED SYSTEM IN SIMULATION ENVIRONMENT 42
- 6.1 VMS Control 42
- 6.2 States for the Q-Learning 43
- 6.3 Actions for the Q-Learning 44
- 6.4 Reward for the Q-Learning 44
- 6.5 Evaluation Measures 45
- 6.6 Implementation in Simulated Environment 47

6.6.1 Look-up table implementation 47
6.6.2 CMAC Implementation 48
6.6.3 Results 48
6.7 Analysis and Results Based on Occupancy States 50

7 FIELD TRIAL 53
7.1 Necessary Steps for Field Implementation of TMS-Can 53
7.2 Evaluation Criteria 55
7.3 Lessons Learned 56

8 SUMMARY AND CONCLUSIONS 59

9 REFERENCES 60

APPENDIX

LIST OF FIGURES

Figure 3-1 System Architecture for TMS-Can 17

Figure 4-1 Map of the study area..... 19

Figure 4-2 Westbound Collector and Express at Yonge St. on Highway 401 (Before Construction)..... 20

Figure 4-3 Westbound Collector and Express at Bayview Ave. on Highway 401 (Before Construction) 21

Figure 4-4 Westbound Collector and Express at Leslie on Highway 401 (Before Construction)..... 21

Figure 4-5 Window Display of Paramics Modeler Paramics Network Coding 24

Figure 4-6 Study Area Network Built in Paramics (West Bound) 25

Figure 4-7 Network Showing the Collector to Express Transfer (West Of Leslie) 25

Figure 4-8 At Yonge Interchange - Collector Reduced to One Lane 26

Figure 4-9 The Standard Reinforcement-Learning Model 27

Figure 4-10 An Agent Behavior in its Environment 28

Figure 4-11 An Example to Illustrate the Concepts of Q-Learning 30

Figure 4-12 Probability of Taking Best Action with Age (E=0.1)..... 33

Figure 4-13 Probability of Taking Best Action with Age (E=0.01)..... 33

Figure 4-14 Cerebellar Model Articulation Controller (CMAC) Tiles 35

Figure 4-15 CMAC Computations 37

Figure 5-1 Corresponding Traffic Zones for The Study Network 38

Figure 6-1 Comparison of Mean Travel Time..... 49

Figure 6-2 Comparison of Mean Speed..... 49

Figure 6-3 Comparison of Mean Travel Time (with Occupancy States) 51

Figure 6-4 Comparison of Mean Speed (with Occupancy States) 52

LIST OF TABLES

Table 2-1 System Functions and sub-functions 6

Table 2-2 Potential Measures of Effectiveness for System Evaluation 7

Table 5-1 Traversal Matrix Generated from EMME/2 Using 2001 TTS data..... 39

Table 5-2 Peak Period Detector Volumes from the Study Area 40

Table 5-3 O-D Matrix for the simulation 41

Table 6-1 Results of Simulation with Q-Learning agents and various Fixed Diversion Percentages 48

Table 6-2 Results for Drivers without Prior Information about Work Zone..... 50

Table 6-3 Result of Simulation with Q-Learning Agents and Various Fixed Diversion Percentages (Using Occupancy States)..... 51

Table 7-1 Detectors Used for Collecting Occupancy Data 54

Table 7-2 Detectors Used for Collecting Outflow 55

GLOSSARY

ADAPTIR	Automated Data Acquisition and Processing of Traffic Information in Real-time
API	Application Programming Interface
AWIS	Automated Work Zone Information Systems
CHIPS	Computerized Highway Information Processing System
CI	Congestion Index
CMAC	Cerebellar Model Articulation Controller
CMS	Changeable Message Signs
CSC	Central System Controller
DSL	Digital Subscribe Line
GUI	Graphical User Input
ICAT	ITS Centre and Testbed at University of Toronto
IRD	International Road Dynamic Inc.
ITS	Intelligent Transportation Systems
LED	Light Emitting Diode
MOE	Measures of Effectiveness
MTO	Ministry of Transportation of Ontario
NMSHTD	New Mexico State Highway and Transportation Department
NTCIP	National Transportation Communications for ITS Protocol
PTMS	Portable Traffic Management System
RL	Reinforcement Learning
RTMS	Remote Traffic Microwave Sensor
TIPS	Travel Time Prediction Systems

TMS-Can	**T**raffic **M**anagement **S**ystem for work zones on **Can**adian roadways
V/C	Volume over Capacity
VMS	Variable Message Sign

1 INTRODUCTION

1.1 Background

Canada's road network is maturing. The increasing growth in traffic and limitations on the expansion of existing infrastructure is taking a toll on the existing transportation network, accelerating the ageing process. Therefore, there is an increasing need for the maintenance and rehabilitation of our roadways. This means more work zones will be created and more traffic congestion as a result of work zones will exist in the road network.

Roadway work zones are inherently less safe and less efficient than areas of roads not under construction. Construction and maintenance activities adversely impact traffic flow by presenting bottlenecks in the roadway system often leading to congestion, delay, and a host of problems ranging from increased vehicle emissions to collisions. Road users surprised by the unexpected presence of a work zone and its accompanying congestion may be unable to change their travel plans in advance, or divert to alternative routes before becoming committed to the affected route. Loss of time, frustration, aggressive driving, and less safe condition may results. It is found that 800 fatality and 37,000 injury accidents in the U.S. every year are related to work zone.

Traditional work zone traffic control focuses on advance warning, site identification and driver guidance. While the controls provide information to motorists, the main purpose of these efforts is to protect workers from traffic, and to allow road users to navigate the work zone safely and efficiently. Over the past several years, many techniques and/or programs have been implemented to enhance work zone safety and facilitate traffic control. These include innovations such as work zone intrusion alarms, portable rumble strips, portable changeable message signs, unmanned radars, solar-assisted arrowboards, flashing stop/slow paddles, and barrier lighting units.

While these measures are an enhancement to traditional traffic control measures, their inability to direct and/or respond to changing traffic conditions, or to invoke any form of dynamic response strategy results in lost opportunities to increase mobility and economic productivity, as well as to reduce costs and environmental impacts.

Therefore, there exists a need to provide accurate and reliable real-time information to drivers as to how long it will take them to clear the work zone, how long they will be delayed by the lane closure, or what speed they can expect ahead of them as they enter the work zone. This information needs to be provided at reasonable intervals in advance of the work zone as well as throughout the work zone. The immediate benefit of this kind of information is that it would help motorists decide whether to stay on the freeway or to take an alternate route since they are now more accurately informed of the expected impact due to the work zone.

Recent research in the development of the Reinforcement Learning (RL) technique, a Machine learning approach, to provide integrated automatic control using Variable Message Signs (VMS) and ramp metering on a freeway/arterial corridor for both recurring and non-recurring congestion, suggests that this powerful approach to corridor and network management can be applicable, albeit in a portable format and at a smaller scale, to temporary conditions traffic management.

1.2 Research Objective

The main goal of this research study is to develop functional and system requirements for an integrated, real-time temporary condition **T**raffic **M**anagement **S**ystem for work zones on **Can**adian roadways (TMS-Can). The research proposes using a control strategy driven by Reinforced Learning (RL) to provide traffic management capabilities and traveler information on traffic routing, and detours to drivers. With real-time information drivers can be aided in route selection allowing them to optimize the available network capacity restrictions imposed due to work zone construction. In other words, the main objective of this research is to provide tools to the road authority to carry out road reconstruction and maintenance activities, while minimizing the disruption of traffic and maximizing the use of available network capacity.

The TMS-Can monitors and assesses changing traffic conditions, invokes traffic management strategies, including alternative routing and demand management, and provides dynamic, real-time information to road users, allowing them to make informed travel decisions. It monitors approaching (upstream) traffic speeds and volume over capacity (V/C), quantifies congestion and determines the best points in time and space to provide information to drivers using advanced traffic control and management strategies.

The TMS-Can uses a combination of electronic sensors, video cameras, software, wireless and cabled communications networks, and electronic signs and variable message signs to manipulate traffic flow. They will be applied to work zones to minimize highway disruptions and increase work zone safety.

1.3 Organization of Report

This research report is organized into ten chapters. The first chapter provides a summary report of the project. Chapter 1 provides a background for work zone related issues and concerns, identifying the motivation behind this research. Chapter 2 provides a detailed summary of the literature reviews on application of Intelligent Transportation Systems on work zone. In Chapter 3, the components of the system will be described. Chapter 4 provides a detailed description of the theoretical background of the proposed reinforcement learning approach. Chapter 5 is devoted to the description travel demand data. Chapter 6 covers the development of the model for the proposed control approach to the study area. The development of the traffic corridor, formulation of the software functions for the application of the proposed control methodology, its implementation,

testing methods and evaluation measures are all described in this chapter. Chapter 7 describes the required steps necessary to implement the concept in a field trial. Chapter 8 provides a conclusion for this project. At the end, Chapter 9 provides a list of references used in this study.

2 LITERATURE REVIEW

Several jurisdictions are using a variety of special devices to monitor traffic conditions in work zones. The equipment in place ranges widely, and includes such use of a wide number of information distribution techniques. Devices such as Variable Message Signs (VMS's), highway advisory radio, traveler information services, the Internet are all deployed. Information is collected in the field using traffic sensors and in some cases information is relayed to as separate traffic control devices.

All of these tools can be integrated together and collectively result in Intelligent Transportation Systems (ITS) that provide information to travelers about work zones. Effectively implemented the tools provide travelers, motorists, transportation agencies, and emergency services with the information necessary to make informed decisions. Specifically, this information can help motorists decide on the selection of the most appropriate travel routes. A transportation agency can use this information to manage its transportation network efficiently. Emergency services can identify incidents in the work zone and determine the most appropriate response.

Overall the literature review found that the basic idea for such systems is to collect the traffic data within the work zone, have this data analyzed and interpreted, and formulate appropriate set of control action send and disseminate control actions to the network and information to travelers in real-time using various output devices.

In this section, a literature review was conducted to investigate existing research on the effectiveness of real- time traveler information systems for the work zones that may be useful in the completion of this assignment. It provides a brief description of relevant available real- time work zone systems. It must be noted that the literature review for this study was conducted in early 2004.

2.1 Specific Systems

2.1.1 Travel Time Prediction System for Freeway Work Zones

The Travel Time Prediction System (TIPS) is a portable system that predicts travel time from any one point to another on a freeway work zone and displays the estimated travel time to drivers via the use of Variable Message Signs (VMS's) that is posted along the freeway. The objective of this system is to provide motorists with travel time information in advance of as well as to those in the work zone on real time basis. This system helps drivers whether to stay on the freeway and transverse through the work zone or take an alternative route. The benefits of the TIPS system are congestion reduction, reduction in driver frustration and increase overall safety in the work zone *(Pant et al., 1998)*.

The system is portable and relies on non-contact sensors, wireless communications, variable message signs, and software capable of predicting travel time. The major

components of this travel time prediction system and their communication links are: (1) Traffic sensors, (2) Intelligent traffic algorithm sited in the on-site (PC), (3) Radios, (4) Cellular phone and modem, (5) Portable variable message signs.

The system works based on three major modules – "Data Collection System", "Data Processing System" and "Information Display System". The system collects traffic flow data such as volume and occupancy every 30-seconds using non-contact sensors. The collected traffic data is transmitted to a personal computer using spread-spectrum wireless radios that are installed at the sensor-stations. The personal computer is located in a trailer at the construction area. A special software within the personal computer processes the collected data by the sensors in real-time and estimates total travel time through the work zone or the travel time between any two points in the work zones. This information is sent from the personal computer to portable variable message signs (VMS's) via cellular phone- modem connection. The software is developed using visual C++ and works on a Windows NT (version 4.0) operating system.

The VMS's have pre-programmed messages in the form of sequences which is input along with messages in the personal computer while setting up the system. The corresponding message is transmitted to the VMS via cellular phone when the personal computer requires the display of the messages on a VMS.

The developed system was tested and validated for a short work zone on westbound I-70 near Huber Heights in Montgomery County, Ohio. The travel time was estimated by using a "three-regime traffic flow model". Model development was done by collection of actual travel time data between sensor stations using video camera. In the next step, the travel times between the locations of the camera were extracted from the videotapes. At the same time, sensor data from the stations were also collected. . Linear and non-linear regression techniques were used to correlate the travel times with the occupancies for estimating travel time.

The TIPS was deployed on a 14-mile segment of northbound Interstate 75 in Dayton, Ohio during the summer and fall of 2000 *(Pant, FHWA)*. The objective of this research was to determine the accuracy of the predicted travel time when compared to the actual travel times. There were total 119 runs recorded by the three crews during the three-day period. The data recorder in each crew also recorded the license plate numbers of private non-commercial vehicles with Ohio license plates. A total of 3177 different license plate numbers were recorded and a questionnaire was sent to each one. A total of 660 completed surveys were returned and analyzed. Linear regression models were used to predict travel times as a function of the actual observed travel times. It was found that about 88% of the actual times recorded for each sign, and for all the signs combined, were within a range of ± 4 minutes of the predicted time. The observed actual times were in range of 8 minutes to 44 minutes. However, a few differences (actual – predicted) as great as 18 minutes were observed. In summary, it was concluded that the real-time TIPS system represents a definite improvement over any static non-real- time display system. It provides in general and most of the time useful and relatively accurate travel time

predictions to the motoring public and appears to be perceived by the motoring public as helpful and useful *(Pant, FHWA)*.

2.1.2 Real-time Advanced Warning and Traffic

Virginia Polytechnic Institute and State University is developing system and functional requirements for a real-time advanced warning and traffic control system for work zone on the I-81 Corridor in Virginia. This system is designed to provide a real-time advanced warning to motorist in order to reduce the impacts of work zone activities on traffic flow and safety. In addition, a set of evaluation criteria and measure of effectiveness was identified to ensure the effectiveness of the system in reducing work zone problems.

The functional requirements developed in this research study include four major functions, namely advanced warning, surveillance, advisory and control. Each of these functions consists of sub functions. The subsystems interact among themselves in order to gather traffic data, analyze it, and based on the analysis, provide appropriate real-time information to the motorists. **Table 2-1** presents the four main functions and their respective sub-functions *(Thommana and Sivanandan, 1997)*.

Table 2-1 System Functions and sub-functions

System Functions	Sub-Functions
Advanced Warning	Congestion Queuing Speed Variation Work Zone Characteristics
Surveillance	Incident Detection Queue Detection Congestion Monitoring Measurement of Traffic flow variables
Advisory	Delay Advisory Speed Advisory Alternate Route Advisory
Control	Change Lanes Speed Limit

In addition to the functional requirements, a system that provides the basis for the type of technology to be employed in assembling the real-time advanced warning and traffic control system is developed. The system requirements developed include real-time operation, credible signage, portability, ease of installation, adaptability, confirmation with the Virginia Work Area Protection Manual standards and guidelines, The National Transportation Communications for ITS Protocol (NTCIP) compliance, open architecture, use of suitable information dissemination tools, cost effectiveness, reliability, and all-weather day/ night operation.

The evaluation criteria are developed to measure the system's capabilities in fulfilling the necessary functional and system requirements. These criteria include functionality, system reliability, performance criteria (measurement accuracy and algorithmic accuracy), installation ease, operation and maintenance ease, technology maturity, adaptability, and cost. As part of the system evaluation, measures of effectiveness (MOEs) for the developed system are identified. The MOEs are categorized according to the following two subdivisions (1) MOEs for the entire system, and (2) MOEs for evaluation of system performance. **Table 2-2** presents the potential MOEs that are identified for each category *(Thommana and Sivanandan, 1997)*.

Table 2-2 Potential Measures of Effectiveness for System Evaluation

Entire System	Reduction in the number of occurrences of congestion and queues
	Reductions in speeds
	Reduction in mean queue length
	Reduction in delays
	Reduction in the number of accidents/severity of accidents
	Speed and advisory compliance
Evaluation of System Performance	Accuracy in measurement of traffic flow variables
	Incident detection rate
	Mean detection time
	False Alarm Rate

2.1.3 Portable Traffic Management System

The Portable Traffic Management System (PTMS) for work zone is a real-time traffic management system that is to provide useful real-time information to motorists about traffic conditions as they approach and pass through a work zone. The basic objective is to improve safety for motorists, improve safety for construction personnel and minimize the work zone related delays experienced by motorists in work zones. The PTMS project is conducted under public-private partnership including: Federal Highway Administration, Minnesota Department of Transportation, and ADDCO, Inc. *(SRF Consulting Group, Inc., 1997)*.

The system is an integration of existing, off-the-shelf, and emerging technologies into a comprehensive portable traffic management system. The PTMS system consists of four basic subsystems: (a) Vehicle Detection/Surveillance (b) Traffic Control Center (c) Driver Information (d) Communications. The vehicle detection/surveillance is used to collect traffic data. It consists of several different components including video camera, portable machine vision devices and speed radars. The video camera are placed at strategic locations in the work zone to record and provide real-time information on traffic data such as traffic volume, speed and incidents. Portable machine devices use video cameras and computers to emulate the function of human eye. These devices gather and record information such speed, volume, occupancy, queue lengths and vehicle

predictions to the motoring public and appears to be perceived by the motoring public as helpful and useful *(Pant, FHWA)*.

2.1.2 Real-time Advanced Warning and Traffic

Virginia Polytechnic Institute and State University is developing system and functional requirements for a real-time advanced warning and traffic control system for work zone on the I-81 Corridor in Virginia. This system is designed to provide a real-time advanced warning to motorist in order to reduce the impacts of work zone activities on traffic flow and safety. In addition, a set of evaluation criteria and measure of effectiveness was identified to ensure the effectiveness of the system in reducing work zone problems.

The functional requirements developed in this research study include four major functions, namely advanced warning, surveillance, advisory and control. Each of these functions consists of sub functions. The subsystems interact among themselves in order to gather traffic data, analyze it, and based on the analysis, provide appropriate real-time information to the motorists. **Table 2-1** presents the four main functions and their respective sub-functions *(Thommana and Sivanandan, 1997)*.

Table 2-1 System Functions and sub-functions

System Functions	Sub-Functions
Advanced Warning	Congestion Queuing Speed Variation Work Zone Characteristics
Surveillance	Incident Detection Queue Detection Congestion Monitoring Measurement of Traffic flow variables
Advisory	Delay Advisory Speed Advisory Alternate Route Advisory
Control	Change Lanes Speed Limit

In addition to the functional requirements, a system that provides the basis for the type of technology to be employed in assembling the real-time advanced warning and traffic control system is developed. The system requirements developed include real-time operation, credible signage, portability, ease of installation, adaptability, confirmation with the Virginia Work Area Protection Manual standards and guidelines, The National Transportation Communications for ITS Protocol (NTCIP) compliance, open architecture, use of suitable information dissemination tools, cost effectiveness, reliability, and all-weather day/ night operation.

The evaluation criteria are developed to measure the system's capabilities in fulfilling the necessary functional and system requirements. These criteria include functionality, system reliability, performance criteria (measurement accuracy and algorithmic accuracy), installation ease, operation and maintenance ease, technology maturity, adaptability, and cost. As part of the system evaluation, measures of effectiveness (MOEs) for the developed system are identified. The MOEs are categorized according to the following two subdivisions (1) MOEs for the entire system, and (2) MOEs for evaluation of system performance. **Table 2-2** presents the potential MOEs that are identified for each category *(Thommana and Sivanandan, 1997).*

Table 2-2 Potential Measures of Effectiveness for System Evaluation

Entire System	Reduction in the number of occurrences of congestion and queues Reductions in speeds Reduction in mean queue length Reduction in delays Reduction in the number of accidents/severity of accidents Speed and advisory compliance
Evaluation of System Performance	Accuracy in measurement of traffic flow variables Incident detection rate Mean detection time False Alarm Rate

2.1.3 Portable Traffic Management System

The Portable Traffic Management System (PTMS) for work zone is a real-time traffic management system that is to provide useful real-time information to motorists about traffic conditions as they approach and pass through a work zone. The basic objective is to improve safety for motorists, improve safety for construction personnel and minimize the work zone related delays experienced by motorists in work zones. The PTMS project is conducted under public-private partnership including: Federal Highway Administration, Minnesota Department of Transportation, and ADDCO, Inc. *(SRF Consulting Group, Inc., 1997).*

The system is an integration of existing, off-the-shelf, and emerging technologies into a comprehensive portable traffic management system. The PTMS system consists of four basic subsystems: (a) Vehicle Detection/Surveillance (b) Traffic Control Center (c) Driver Information (d) Communications. The vehicle detection/surveillance is used to collect traffic data. It consists of several different components including video camera, portable machine vision devices and speed radars. The video camera are placed at strategic locations in the work zone to record and provide real-time information on traffic data such as traffic volume, speed and incidents. Portable machine devices use video cameras and computers to emulate the function of human eye. These devices gather and record information such speed, volume, occupancy, queue lengths and vehicle

classification in the work zones. The speed radars are used to collect speed data. The speed radars are placed on the locations in the work zone that are not equipped with machine vision in order to provide additional information. The collected data from the vehicle detection/surveillance subsystem is transmitted to the traffic control center at the Minnesota/DOT Traffic Management Center via ISDN link. The data is reviewed and analyzed by system operators and decisions are made regarding the changes needed to improve traffic flow through the work zone. The traffic control changes are made by relaying messages to the motorists via the driver information subsystem, which consists of full-size portable Variable Message Signs (VMS's) and smaller work zone portable variable message signs. This information is also made available to the public on a World Wide Web page via the Internet. The communication subsystem relies on spread spectrum radio, cellular links, and ISDN. Each of these communication devices is used for specific links within the system. The guidance subsystem is a moving light setup on a portable concrete barrier. This will be used in areas where lanes are tapering in order to enhance lane visibility and improve the motorist's visual comprehension of the roadway geometry. This subsystem will also include a master controller and a radio link to the traffic control center.

The operational test of the PTMS system took place on Interstate 94 between downtown Minneapolis and downtown St. Paul, Minnesota and on Interstate 35 in the southern portion of the Twin Cites metropolitan area. The evaluation process for this project consisted of four parts:

- Analysis of Traffic Impacts
- Analysis of System Operation
- Analysis of User Reaction
- Analysis of System Costs

In the following subsections the summary results of analysis for each evaluation process was posted by SRF Consulting Group, Inc., 1997.

2.1.3.1 Analysis of Traffic Impacts

- There was a significant increase in the traffic volume that moved through the work zone when the PTMS was in operation (3.6 percent higher in the morning peak period and 6.6 percent higher in the afternoon peak period). The more orderly flow of traffic appears to have resulted in increased capacity.
- There was a significant decrease in the traffic volume that exited I-94 to TH 280 in the afternoon peak period (5.3 percent lower). This is likely due to increased driver confidence in traveling through the work zone because of the real-time traffic information provided by the PTMS.

- The PTMS decreased the variability in speed for traffic traveling within the work zone by over 70 percent, which suggests improved safety.

- The PTMS decreased the average speed for traffic approaching the work zone by 9 miles per hour, which also suggests an improvement in safety as vehicles approaching the work zone slowed down sooner.

2.1.3.2 Analysis of System Operation

- PTMS can be deployed in a wide variety of work zones with relative ease.

- The actual time to prepare, deploy and set up each node is relatively short, especially given the complexity of the PTMS equipment and its capabilities.

- The PTMS is a unit that would be set up in work zones with special traffic needs.

- The PTMS was successfully installed and operated at the TMC.

- The PTMS was developed using TCP/IP communications protocols to simplify integration into existing TMCs.

- The overall system operation was successful.

- The machine vision cameras mounted high on the towers require care that construction activities do not inadvertently cause the skids to move.

- The Autoscope setup is not complex but requires some special training.

- The PTMS video and data can be successfully transmitted over spread spectrum radios.

- Care must be taken in setting the nodes to avoid multi-path interference problems.

2.1.3.3 Analysis of User Reaction

- 66 percent of the drivers surveyed remembered seeing the lighted PTMS VMS messages.

- Of those remembering the PTMS VMS messages, most remembered specific messages and felt more informed about traffic conditions.

- Of the specific messages remembered, very high percentages were considered easy to read, easy to understand, useful and to correctly reflect traffic conditions.

- System operators indicated that the PTMS could be mastered in 2 to 3 hours was easy to operate and was reliable.

- Transit operators said the video surveillance feature would be very useful for rerouting buses from work zones with significant delays.

- Focus groups called for real-time traffic information that was easy to read and the motorist survey indicated the PTMS was successful in these respects.

2.1.3.4 Analysis of System Costs

- The cost of an operational PTMS depends on the number of PTMS nodes needed and the additional optional equipment needed.

- The cost for a basic PTMS with all necessary communications equipment is $78,850.

- Additional basic nodes would cost $59,850.

- Deployment cost will vary based on PTMS delivery distance, number of PTMS nodes and ISDN installation cost and monthly charges

- The feature being developed to automatically change sign messages based on speeds detected will significantly reduce staff time to operate the system.

2.1.4 Automated Work Zone Information Systems

The Arkansas Sate Highway and Transportation Department in an effort to improve safety and manage congestion in work zone has deployed Automated Work Zone Information Systems (AWIS). The main purposes of the developed systems were as follows *(Tudor et al., 2003)*:

- Prevent/Reduce Rear-End and other crashes

- Improve Work Site Safety

- Provide Real-time Information to the Motorist Regarding Potential Backups caused by Lane Closures

- Improve Incident Management

- Improve Congestion Management

The AWIS installed on interstate rehabilitation projects in Arkansas. The system used are ADAPTIR (Automated Data Acquisition and Processing of Traffic Information in Real-time) by Scientex, Crop. and CHIPS (Computerized Highway Information Processing System) by ASTI Transportation Systems. The characteristics of each system are described in following sections.

2.1.4.1 "ADAPTIR" System

This system consists of a Central System Controller, two highway advisory radios, five traffic sensors, five changeable message signs and two supplemental speed stations per lane closure. The system provides real-time information regarding current traffic status to motorists and forewarning them of congestion. The system also provides alternate route messages in the event of congestion. The system communicates by landline telephone to the highway advisory radio system. The system is designed for the VMS to display downstream traffic speed information followed by delay information with preset scenarios. If the speed differential between traffic sensors exceeds 10 mph, the upstream VMS will display "REDUCE SPEEDTO XX MPH", followed by the message "YY MINUTE DELAY". If the speed differential is below 10 mph, the VMS will only display the delay information message. The Central System Controller is programmed for a 600 second cycle time *(Tudor et al., 2003)*.

The highway advisory radio consists of three separate phases or blocks that are combined into one message. The first message phase is an introductory message providing general project information. The second phase is a daily or weekly message explaining the lane closures planned for the day or week. The third phase is the message that fits the real-time situation in effect at the lane closure.

The ADATIR system was installed in the interstate 40 from Highway 31 to the city limits of Carlisle during periods of May 2000 and November 2001. Four problems were encountered: (1) Delay estimations were not accurate enough. Therefore, the time delay messages were simplified to report "EXPECT DELAYS" or "EXPECT LONG DELAYS" (2) Remote Traffic Microwave Sensor (RTMS) units were initially used for the traffic sensors. However, it was found that the RTMS units are hard to calibrate and operate due to high percentage of trucks. Therefore, Doppler radar units were used instead of the RTMS units (3) It was determined the cycle length of 600 seconds was too long for optimum accuracy. Therefore, the cycle length was changed to 263 seconds (4) The Central System Controller was unable to contact the highway advisory radio via landline telephone in some cases.

In order to determine the accuracy of the system, the projected travel time was compared to with the actual travel time. The acceptable range of error for the difference in time between the actual and projected travel times was defined as less than or equal to five

minutes. It was found that the system both overestimated and underestimated the actual delay. The maximum overestimate for delay was 12 minutes 35. On the other hand, the maximum underestimated for delay was 2 hours 11minutes 10 seconds.

In order to examine the system effectiveness, a fatal and rear-end crash analysis was conducted which compared two locations on Interstate 40 with the site with ADAPTIR system. It was concluded that a reduction in fatal crash rate on the site with the ADAPTIR system. A mix results were achieved when the rare-end crash rates were analyzed.

2.1.4.2 "CHIPS" System

The CHIPS system consists of a Central System Controller, a host computer, a VMS and six traffic sensors. The system also communicates by landline telephone to a highway advisory radio system. The system provides motorist with real-time traffic, delay and diversion advisories.

The data acquisition is performed using traffic sensors located along the entire length of work zone. The Central System Controller (CSC) located in the contractor's office retrieves this acquired data. The controller is programmed for a 350 second cycle time that includes communication from the traffic sensors to the VMS, highway advisory radio, a host computer in the Engineer's office, pagers and email transmissions using 80 predetermined scenarios. In the event of deteriorating traffic conditions, the pagers and emails are transmitted to warn drivers of delays, low traffic speeds, and other hazardous conditions. The predetermined scenarios include: Slow Traffic Ahead- Be Prepared To Stop, XX Mile Backup YY Miles Ahead- Be Prepared To Stop, Drive Safely-Buckle Up.

The system was implemented in the interstate 40 from Highway 67 to the Lonke County Line, in Pulaski County. This 8.6 mile urban segment had an average daily traffic of 44,000 vehicles per day with a truck percent of 35% for year 2001. There were four problems encountered upon operation of the system. They were as follows: (1) Phone lines for the Internet and software support were lost and could no be accessed. The contractor considered using Digital Subscribe Line (DSL) or cable modem to remedy this problem (2) During periods of heavy cellular phone usage, communication with the highway advisory radio via cell phone experienced cell carrier drops in service. This problem was addressed by changing to POTS (Plain Old Telephone System) lines. (3) Touch tone pagers did not convey accurate enough information, therefore it was switched to Arch wireless pagers. These pagers send text messages to the appropriate personnel with more accurate real-time information (4) Additional computer was added to the Central System Controller in order to enhance the performance of the system. The additional computer was used for handling such tasks as emailing, sending pages and updating web pages, while the "master" computer continued to communicate with the traffic sensors and VMSs.

In order to determine the accuracy of the system, the messages on the VMS and the highway advisor radio compared with the actual traffic condition. The starting point for the data collection was at the location of the first visible VMS. The terminal point was at the end of the work zone. It was found that the 90 percent of messages matched the actual conditions.

2.1.5 Condition-Responsive Work Zone Traffic Control System

The Condition-Responsive Work Zone Traffic Control system is a real-time work zone system developed by the Scientex Corporation for the Maryland State Highway Administration in cooperation with the FHWA. The system monitors traffic flow at a number of locations in advance of the work zone. Based on traffic conditions, the system uses CMSs and highway advisory radio to advise drivers about operating speeds, lane closures, and congestion ahead and to recommend diversion to alternate routes. The basic components of this system are: (1) Changeable Message Signs (CMSs) (2) Traffic Sensor (3) Highway Advisory Radio (4) Trailer with Solar Power System (5) Radio Remote Station (6) Central System Controller *(McCoy et al, 1999).*

2.1.6 Intelligent Transportation Systems in New Mexico

The New Mexico State Highway and Transportation Department (NMSHTD) used ITS to reconstruct "The Big I" interchange, where I-40 and I-20 intersect in Albuquerque. The primary goal of "The Big I" project was to provide efficient incident management with additional goals of providing traffic management capabilities and traveler information on traffic routing, detours. Through the system of cameras, portable DMS trailers, portable traffic management system, and highway advisory radio units NMSHTD was able quickly identify incidents, determine an appropriate response and reduce the average time to respond and clear an incident from 45 minutes to 25 minutes *(FHWA, November 2002).*

The data on traffic conditions was acquired via roadside camera and transmitted to a central base station sever in the traffic management center. The base station server processed and displayed the images collected from camera. Any problems such as incidents were identified by NMSHDT staff person who is monitoring the segment of highway section manually. The central base station server displayed messages about detour routes and major incidents. The messages could be activated for display either manually or automatically. The traffic condition information was also available via the Internet. The system also archived camera imagery and created electronic incident reports.

2.2 Other ITS Work Zone Systems

The systems mentioned in previous section are used to provide real-time information within work zone for travelers. In addition to the system highlighted previously, there are many other systems that employ other applications or technique of ITS in work zones. This section provides a brief summary of other ITS applications and advanced technology approaches aimed at improving safety, speed management system, and capacity.

It must be emphasized that the systems described in this section are not directly related to the system that will be implemented in this particular research but it provides insight for different strategies used in work zone by applying ITS.

2.2.1 Dynamic Work Zone Safety System

The Dynamic Work Zone Safety System developed by International Road Dynamic Inc. (IRD) in partnership with Indiana Department of Transportation addresses the specific problem of merging traffic from two lanes in to a single lane. The system consists of the following components: non-intrusive traffic sensors, interface controllers, communication devices, regulatory signboard with flashers and trailer, and solar power equipment and batteries *(Klashinsky and Bushman, 2000).*

The system is designed to prevent dangerous merging in the tapered approach to work zones by creating a dynamic no-passing zone. The dynamic no passing zone is created by using a series of traffic signs and traffic signs with flashing beacons such as a "DO NOT PASS" sign and "WHEN FLASHING" sign. As sensors detect that traffic is backing up, the next upstream sign begins to flash. The length of the no-passing zone depends on the length of the traffic back up. As congestion occurs at a certain point, the no passing zone must be moved upstream of this point to provide vehicles with time and space to merge while traffic is still moving and there are sufficient opening *(Klashinsky and Bushman, 2000).*

2.2.2 SpeedGuard System

The SpeedGuard system is a portable system designed to improve safety and efficiency of traffic operations by informing drivers of their speed. The objective of the system is to reduce both the average speed of vehicles and speed variations in work zones. The system consists of a portable trailer unit that uses radar to measure vehicle speed and a panel with 61cm light-emitting diode (LED) numerals on which vehicle speeds are displayed. The system is bulletproof and is camera-ready to allow photo enforcement *(FHWA, September 2002).*

The system was deployed on a rural section of I-70 in Kansas and on I-80 in rural section of Nebraska. The system was evaluated by using a before-and-after study approach. It was found that a significant reduction in mean speeds, 85th percentile speed, the

percentage of drivers exceeding the posted limit, and speed variation in Kansas. In Nebraska, evaluation team found that the system to be effective in reducing speed by about 8 km/h and increasing the uniformity of speeds.

2.2.3 Wizard CB Alert System

The Wizard CB Alert System is used to automatically broadcast advisory messages to warn drivers of traffic or road conditions/hazards ahead and is primarily targeted at truckers. The system broadcasts a pre-recorded message of approximately 7 to 10 seconds duration warning drivers over a citizen band channel. The message contains information about downstream construction activates or other conditions associated with reduced traffic speeds. The objective of the system is to give advance warning of upcoming delays to the drivers of heavy vehicles to enable them safely *(Maze et al, 2000).*

The system was evaluated in Iowa in July 1999 during a striping operation on a rural section of I-35. It was found that 75 percent of truckers surveyed had heard the message. All of the drivers who reported hearing the message thought this approach were worthwhile; 89 percent of the drivers thought the message to be effective.

2.2.4 Work Zone Intrusion Alarms

The work zone intrusion is a device developed to enhance the safety of workers by giving them an audible warning in the event of work area intrusion. The system detects vehicles entering the buffer area between work crews in the workspace and vehicle driving past the work zone, and provides warning to alert workers. The alarm provides workers with an estimated four to seven seconds warning to avoid the intruder. The system uses various technologies such as infrared, ultrasonic, microwave, or pneumatic tubes, to detect the intruding vehicle *(Canadian Strategic Highway Research Program, 1995).*

3 SYSTEM REQUIREMENTS FOR TMS-Can

As highlighted in previous sections, the main focus of this project is on the application of Intelligent Transportation Systems to work zone traffic management on highways by developing a system which provides real-time information to road users, allowing them to make informed travel decisions.

The TMS-Can, like most real time advanced warning and traffic control system, consists of two components, namely software and hardware components, and several subsystems that are integrated together to form the system. The subsystems work together in collecting data and analyzing it, and providing appropriate real time information to the travelers.

3.1 Software Component

The software component of such a system is composed of two main subcomponents:

1. A sub-system that tracks traffic measurements upstream of the construction site on the highway and quantifies congestion. The purpose of this subsystem is to determine the real-time level of service of traffic and, using this input, determine at which point dynamic traffic control should be enacted. The concept of "Congestion Index" (CI) will be used for accomplishing this. CI is the ratio of the actual speed to a reasonable reference speed. It will be investigated during this research that at what index, control actions should be triggered.

2. A sub-system that formulates an integrated set of control actions using the primary tool: a variable message sign. As appropriate any available ramp metering systems existing upstream of the construction zone could also be utilized. The system uses Reinforcement Learning (RL) method, an Artificial Intelligence method for machine learning, to provide a single integrated optimal control agent for a freeway-arterial corridor in the construction zone. A Q-learning model, which uses the Reinforcement Learning method, is used on a microscopic simulation model using Paramics micro-simulation program *(PARAllel MICroscopic Simulation) (Quadstone Ltd, 2000)* to develop this tool.

3.2 Hardware Component

The basic components of the system are:

1. *Electronic sensors:* If loops or other detectors are available in the construction zone, the system would use the existing detectors. Otherwise, in the case of roadway without available loops, the data acquisition is performed using non-

pavement intrusive sensors such as road-tubes or remote detectors within and upstream of a work zone.

2. *Output devices:* such as Variable Message Signs (VMS) used to display the message.

3. *Wireless and Cabled Communications Networks:* Wireless and Cabled Communications Networks are used to the collected information from the software to the control center and vise versa. The cabled communication links will be used for the area, which facilitates the use of the system in areas not yet supported by cellular service or during periods of heavy cellular phone usage.

3.3 TMS-Can Architecture

The TMS-Can works as follows. Traffic flow information such as speed and occupancy for upstream and downstream of the work zone is collected by electronic sensors. The collected data is sent to the traffic control center at University of Toronto ITS center where decisions regarding the status of traffic within the work zone are made. After reviewing and analyzing the data, the traffic data center sends out appropriate advanced warning and advisory messages to the drivers using the output devices via wireless communications networks. **Figure 3-1** shows system architecture for TMS-Can.

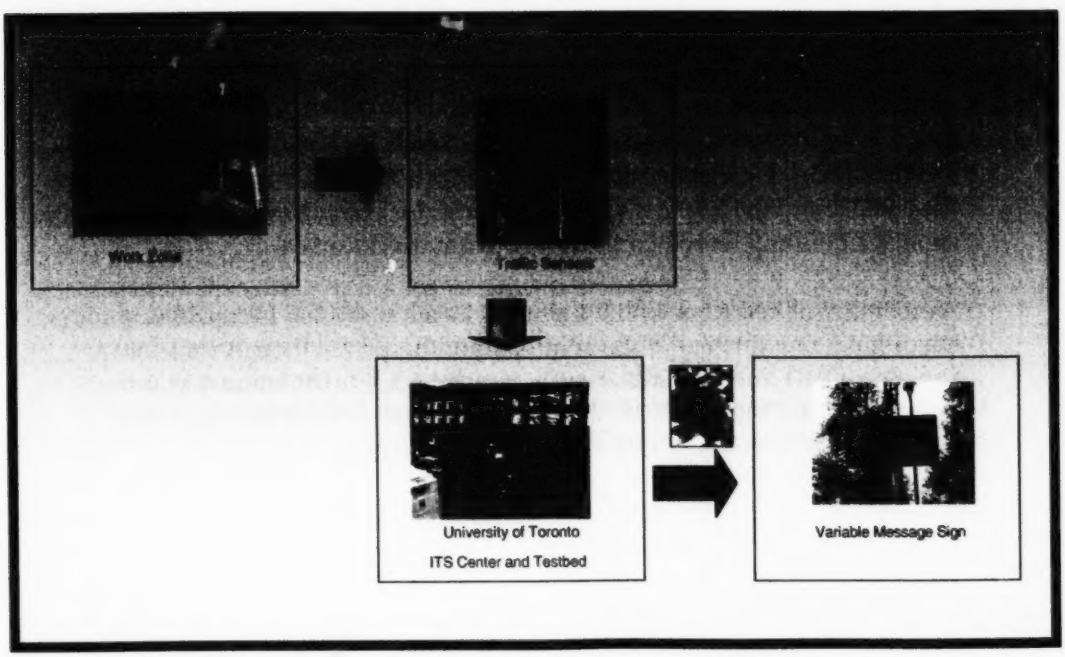


Figure 3-1 System Architecture for TMS-Can

pavement intrusive sensors such as road-tubes or remote detectors within and upstream of a work zone.

2. *Output devices:* such as Variable Message Signs (VMS) used to display the message.

3. *Wireless and Cabled Communications Networks:* Wireless and Cabled Communications Networks are used to the collected information from the software to the control center and vise versa. The cabled communication links will be used for the area, which facilitates the use of the system in areas not yet supported by cellular service or during periods of heavy cellular phone usage.

3.3 TMS-Can Architecture

The TMS-Can works as follows. Traffic flow information such as speed and occupancy for upstream and downstream of the work zone is collected by electronic sensors. The collected data is sent to the traffic control center at University of Toronto ITS center where decisions regarding the status of traffic within the work zone are made. After reviewing and analyzing the data, the traffic data center sends out appropriate advanced warning and advisory messages to the drivers using the output devices via wireless communications networks. **Figure 3-1** shows system architecture for TMS-Can.

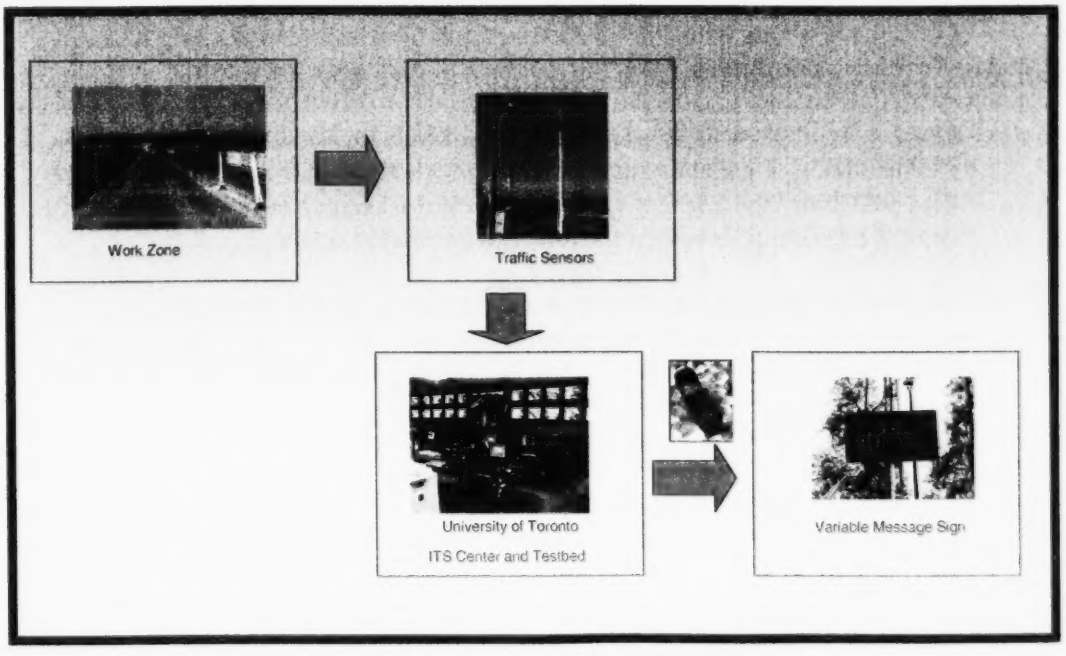


Figure 3-1 System Architecture for TMS-Can

4 SYSTEM FUNCTIONALITY

The TMS-Can is comprised of several components that work together to operate as a single system. To develop such a system, an artificial intelligence based computer agent is used to provide real time traffic control on work zones. It focuses on providing real time messages through VMS (Variable Message Signs) in order to direct the traffic effectively to minimize system wide delay and congestion. Reinforcement Learning Technique from the Machine Learning community is utilized to develop the agent. Reinforcement Learning is an approach whereby the control agent directly learns to map sensed system states to optimal actions. A micro simulation tool – Paramics *(Quadstone Ltd, 2000)* has been used to train the agent in an offline mode in a simulated environment. The software development to incorporate the Reinforcement Learning agent has been made as a plugin using Paramics Application Programming Interface (API).

4.1 Study Area

The Ministry of Transportation of Ontario (MTO) allotted a construction site that could be used for this project and provided the necessary access to information and data required for the development of the network, its simulation with existing traffic conditions and testing. The study area selected for the work-zone project was the construction site on the westbound collector of Highway 401, in Toronto, west of Yonge St. at Hogg's Hollow. **Figure 4-1** shows the map of the study area. Construction had reduced the collector to one lane at the site, which greatly affects the collector capacity. The site selected was anticipated to be most suitable for the implementation of the proposed control algorithm based on the presence of the following features:

- There is a Collector to Express transfer (between Bayview Ave. and Leslie St.) well upstream of the congested location, which allows for diverting traffic from the Collector to the Express via VMS (the major control measure adopted in the control algorithm), based on the downstream traffic situation.

- Availability of detectors within the study location –which is considered as the major input requirement of the control algorithm. Even though some of the permanent detectors within the study area were not functioning due to the construction activities, the MTO has replaced some of the detectors with temporary RTMS (Remote Traffic Microwave Sensor).

- An Express to Collector transfer, west of the construction zone between Yonge St. and Avenue Rd. This delineates the study area from the upstream transfer to the downstream transfer, which includes the congested work-zone. Traffic that is diverted to the Express may get an opportunity to come back to the Collector past the congested zone, if required.

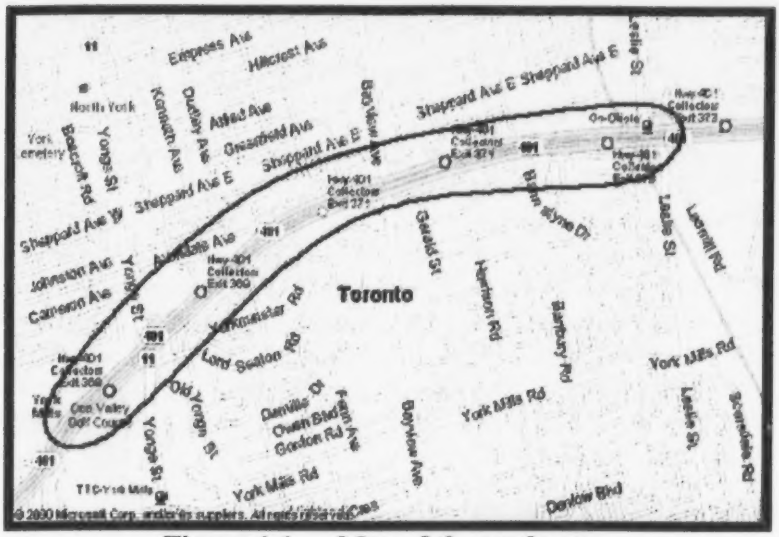


Figure 4-1 Map of the study area

4.2 Field Visit

A site visit was conducted to obtain a better understanding of the character of the freeway and to observe specific characteristics. Included in **Appendix A** is a photograph log of our site visit. It was observed that the Collector was reduced to 1 lane from 3 lanes at the Yonge interchange due to the construction. The Collector between Bayview Avenue and Yonge Street reduced to two lanes. This creates a bottleneck that will greatly affect the Collector capacity. During the site visit, a collision occurred on the Expressway that caused both Collector and Express to experience a major traffic jam.

4.3 Software Development

The part of the Highway 401 study area from Leslie St. to Avenue Rd. including both Collector and Express along with its on and off-ramps has been built using the Paramics simulator. The network geometry has been coded using the base map of the area as an overlay file. Necessary modifications were made to reflect the current geometry during construction. Demand matrices for the area as obtained from the EMME/2 *(INRO Consultants INC, 1998)* were used for the preliminary model building. These Origin and Destination (O-D) volumes were adjusted using the existing link flows. The existing link volumes during the morning one-hour peak period (8:00 a.m. to 9:00 a.m.) were collected from the appropriate detectors on the study area.

As highlighted at the beginning of this section, Reinforcement Learning Technique is used for training the computer agent in this research. The main input to the learning agent is the information about the state of the network as obtained from the detectors on the study area like speed, volume and occupancy. The agent was trained in Parmaics

environment for a few weeks. The Reinforcement Learning agent receives the information about the traffic (speed, volume, occupancy) on the Collector and the Express as well as the incoming and off going traffic from ramps. On receiving the state of the network, the agent provides the optimal routing strategy as percentage number of vehicles to be rerouted from Collector along the Express at every five minutes.

Detectors on both Collector and Express as well as on the ramps should be in working condition to extract real time data. This is one of the major hurdles for the development and implementation of the project as most of the detectors at the work zone area may be dead or unreliable due to the construction activities. In the present work site also, only few of the detectors are reliable and some of them were replaced with the RTMS (Remote Traffic Microwave Sensor) the accuracy of which is uncertain. **Figures 4-2 to 4-4** show schematic diagram of the intersection at Yonge St. and Bayview Avenue and Leslie respectively before construction.

Figure 4-2 Westbound Collector and Express at Yonge St. on Highway 401 (Before Construction)

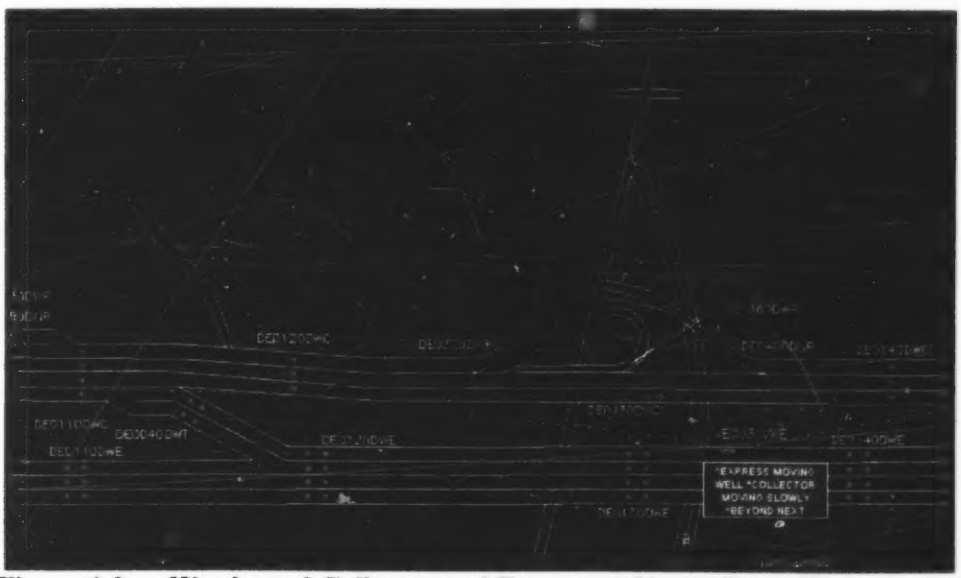
Figure 4-3 Westbound Collector and Express at Bayview Ave. on Highway 401 (Before Construction)

Figure 4-4 Westbound Collector and Express at Leslie on Highway 401 (Before Construction)

The simulated experiment with the learning agent conducted for the study area for the morning peak period with the existing traffic conditions has shown considerable improvement over the base case. The network status with the current diversion percentage was considered as the base case.

environment for a few weeks. The Reinforcement Learning agent receives the information about the traffic (speed, volume, occupancy) on the Collector and the Express as well as the incoming and off going traffic from ramps. On receiving the state of the network, the agent provides the optimal routing strategy as percentage number of vehicles to be rerouted from Collector along the Express at every five minutes.

Detectors on both Collector and Express as well as on the ramps should be in working condition to extract real time data. This is one of the major hurdles for the development and implementation of the project as most of the detectors at the work zone area may be dead or unreliable due to the construction activities. In the present work site also, only few of the detectors are reliable and some of them were replaced with the RTMS (Remote Traffic Microwave Sensor) the accuracy of which is uncertain. **Figures 4-2 to 4-4** show schematic diagram of the intersection at Yonge St. and Bayview Avenue and Leslie respectively before construction.


DET1400WE
DET3040WT
DET1400WE
DET1400WE
DET1400WE

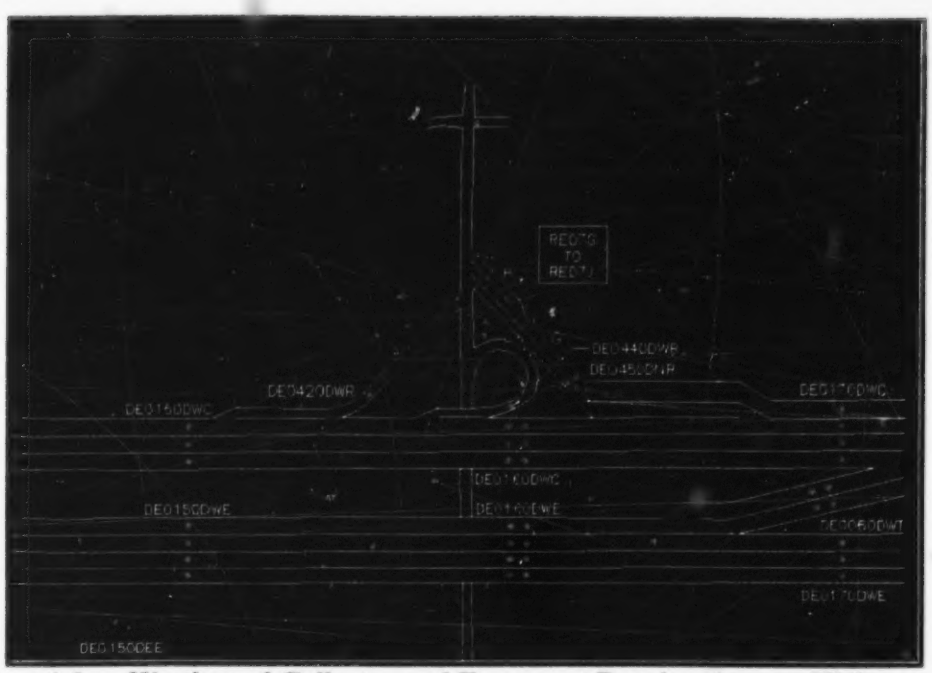


Figure 4-3 Westbound Collector and Express at Bayview Ave. on Highway 401 (Before Construction)

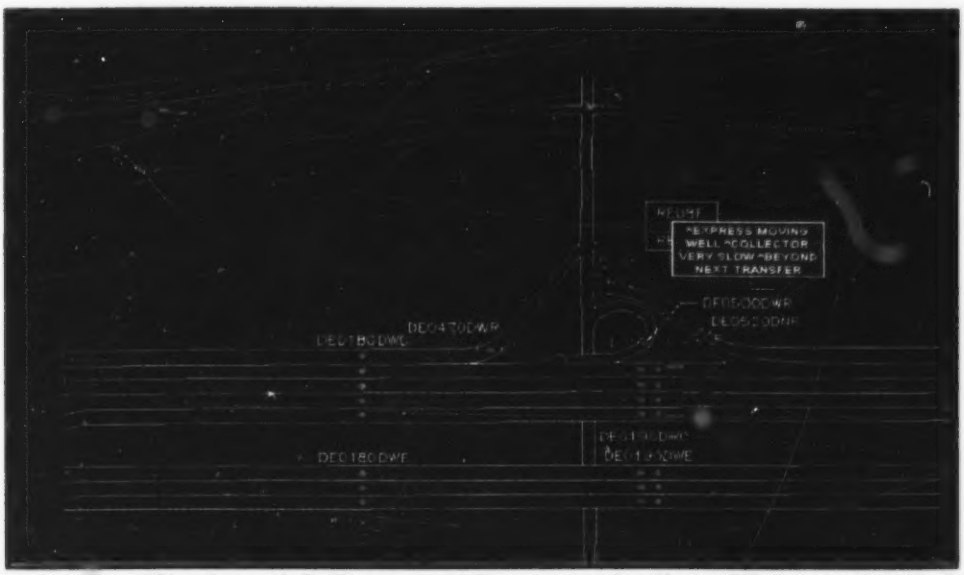


Figure 4-4 Westbound Collector and Express at Leslie on Highway 401 (Before Construction)

The simulated experiment with the learning agent conducted for the study area for the morning peak period with the existing traffic conditions has shown considerable improvement over the base case. The network status with the current diversion percentage was considered as the base case.

The agent, which is already trained in the simulated environment, is now ready for field implementation, wherein the detectors at the study area will provide the current state on the network and with this information the agent will provide the optimal percentage of traffic to be 'allowed' on the freeway through the congested area. This information is to be effected by casting appropriate message signs on the VMS.

4.4 The Reinforcement Learning Based Approach

This section describes some the theoretical background for Paramics simulator. Also this chapter explores the theoretical background of reinforcement learning, its main elements such as its representation, exploration vs. exploitation, learning rate, etc. A detailed account is provided on the Q-learning algorithm. Finally, the requirement of the function approximation technique with reinforcement learning is discussed and the approach selected for this research– Cerebeller Model Articulation Controller (CMAC) is presented.

4.4.1 Paramics Simulator

Traffic simulation programs are useful tools for modeling and predicting traffic flow, evaluating traffic management strategies, and designing roadway facilities before field operations. This project employs a microscopic traffic simulation program, Paramics, to implement the reinforcement Learning algorithm and to train the agent in an offline mode. Paramics, developed by Quadstone Limited, is a suite of high performance tools for microscopic simulation, consisting of Paramics Modeler, Processor, Analyzer, Programmer, and Monitor *(Quadstone Ltd, 2000).*

Paramics comprises a suite of high performance, user programmable software for microscopic traffic simulation. It enables finer descriptions of network structure, vehicle types, vehicle performance, and various kinds of traffic flow information acquisition. The most valuable feature of Paramics is its Application Programming Interface (API). It enables the user to override the default functions in Paramics, such as its car following, merging, vehicle release, and route choice codes. Users can also add their own APIs to implement many traffic control/management strategies within Paramics.

The Paramics version 3.0 by Quadstone Limited consists of five software modules: Modeller, Processor, Analyzer, Programmer, and Monitor *(Paramics Modeller V 3.0 User Guide).* Brief descriptions of each software modules are as follow:

1. Modeler is the main module of the Paramics suite. It provides three fundamental operations: geographic and traffic data input, simulation, and acquisition of statistical traffic data output using both graphic user interface and text coding.

2. Processor sets up and runs the traffic simulation in batch mode without visualizing the model. Processor is useful when running sets of different simulations in much shorter duration.

3. Analyzer is an analysis tool for displaying the output from simulation. Analyzer can provide a range of statistics such as simulated vehicle paths, traffic flow volume by link and turn, maximum queue lengths, and simulated journey times.

4. Programmer is a framework that allows customization of many default Paramics features, such as input/output of traffic data, the specification of driver and vehicle behavior, modification of routing and assignment algorithms. Access to this module is provided through the Application Programming Interface (API).

5. Monitor estimates the levels of traffic emission pollution on a road network.

4.4.1.1 Paramics Model Development Cycle

The Paramics development cycle is summarised in the following seven steps *(Millar, 1999)*:

1. Creating a new Paramics network and embedding template road geometry file.

2. Building a road network by adding nodes, links and zones and coding detailed lane and junction descriptions.

3. Building demand matrices from origin/destination data and including fixed demand data such as bus routes.

4. Assigning traffic using an appropriate assignment technique.

5. Collecting and analysing model results.

6. Calibrating base conditions by comparing model results to observed data.

7. Validate the base model against independent data.

The traffic assignment stage can use all-or-nothing techniques, stochastic assignment, dynamic feedback assignment or combinations of these. Vehicles in Paramics do not carry routing information for more than two links in advance, other than knowing which destination they are headed for. As they progress through a network they look at the cost of alternative routes and choose the lowest cost route as they see it at that time. This gives vehicles the flexibility to re-route due to network changes e.g. lane closures or to incidents or events as they happen. The only exceptions are fixed route vehicles, such as buses, which are coded to run along predefined routes.

Figure 4-5 depicts the main display windows of Modeler. Most of the Modeler options can be selected by the pull down menus located in the top of the main window. Network can be constructed by using the editor button located in the left bottom toolbars or by coding nodes, links, and link characteristics in the text file format. The main window also contains a number of other sub-windows such as parameter variation sliders, the dashboard view window, and the interactive traffic data window.

Paramics Modeler and Programmer serve as the major modules for this project.

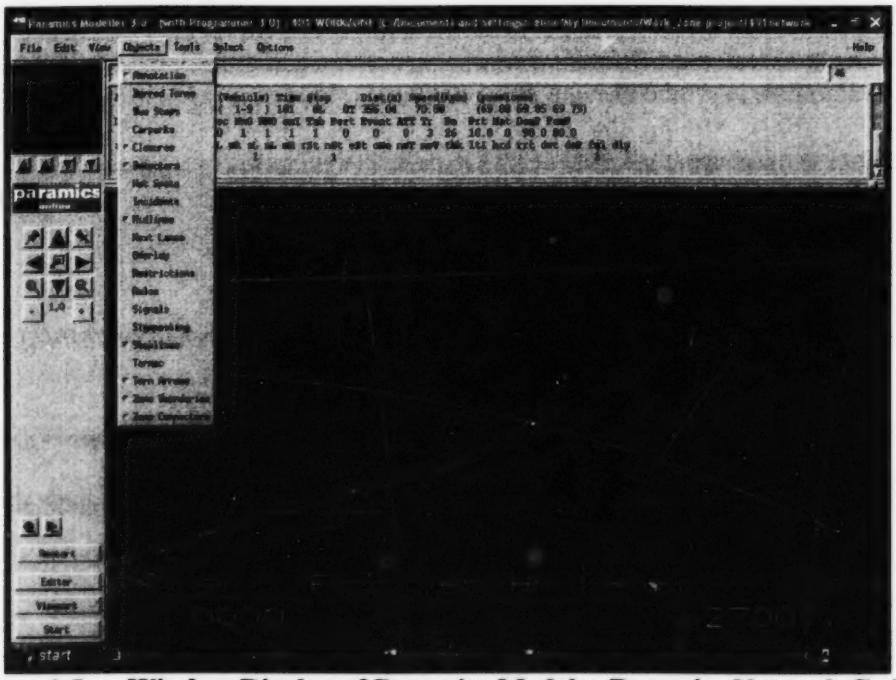


Figure 4-5 Window Display of Paramics Modeler Paramics Network Coding

4.4.1.2 Paramics Network Building

Modeler requires two main inputs viz. road network data and travel demand data. Paramics simulation starts with the building up of the study area network. Detailed road layout plans can be read directly into Paramics and used as a template to build the model network. Overlays can be used as a template to build a network model. These can be read into the model as an AutoCAD (DXF) file. Nodes, links, zones, ramps, detectors, turning movements, junctions, VMS are all coded as per the geometry and information either by GUI (Graphical User Input) or by text files. **Figures 4-6 to 4-8** show the study area network built in Paramics.

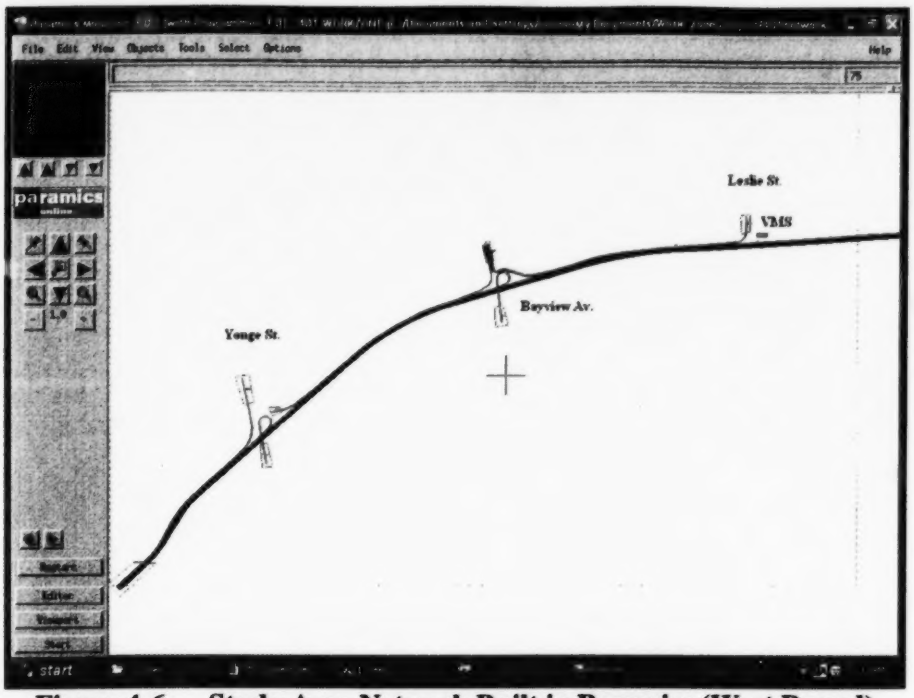


Figure 4-6 Study Area Network Built in Paramics (West Bound)

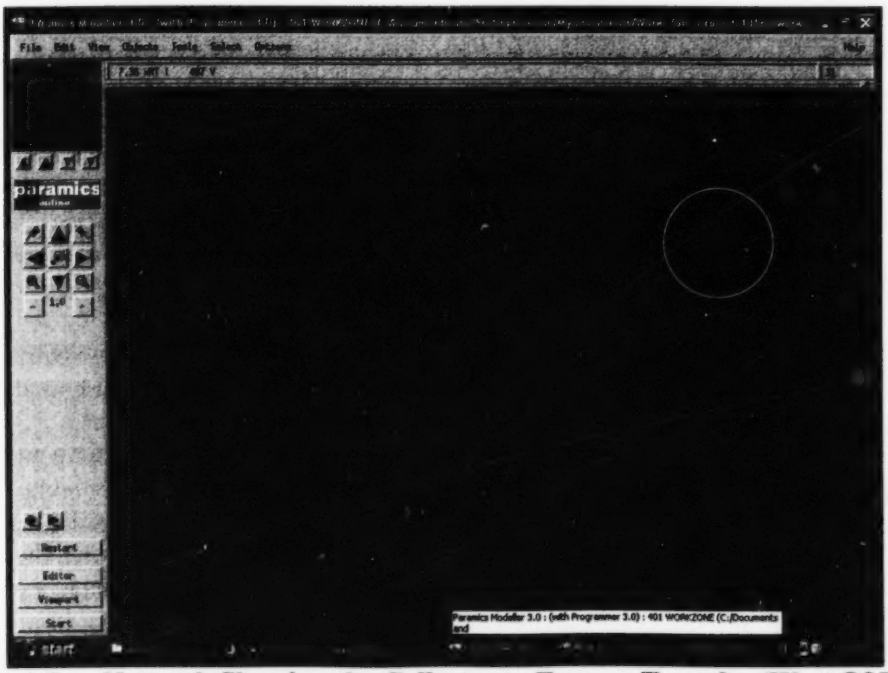


Figure 4-7 Network Showing the Collector to Express Transfer (West Of Leslie)

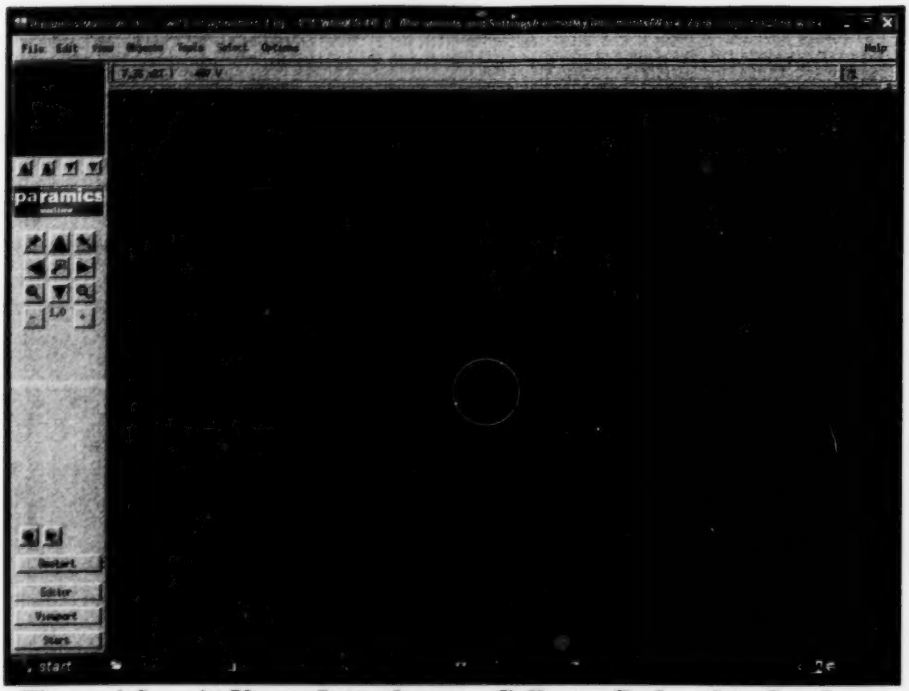


Figure 4-8 At Yonge Interchange - Collector Reduced to One Lane

Travel demand in Paramics is defined by a matrix of origin to destination trips, specified in the demands file. Trips are proportioned into vehicle types and are profiled by 5 minute time periods for a specified duration (1 hour peak). Trips are released from and arrived at geographical areas called zones. The matrix of trips to and from zones is defined in demands file directly. The vehicle file defines the characteristics of each vehicle. Detectors in Paramics work in the same way as the real induction loop detectors, but can also gather some traffic data that could not be collected by real detectors. The point data like speed, occupancy, headway, flow rate, count etc are recorded every time a vehicle passes over a detector.

Once the network is ready as per the road geometry and the demand matrix, vehicles, specifications all provided, still a major share of the work is necessary to fine-tune the network to simulate the traffic behavior close to the reality. One has to run the simulations several times adjusting the parameters like cost factors, reaction time and mean headway. Similarly, much care should be taken to adjust the geometry coding elements like sign posting distance, stop lines and kerb points. These will significantly affect the speed of the vehicle during simulation run. Paramics provides signposting function to assign the location of a signpost at which drivers are aware of the hazard (e.g. road narrowing, restrictions, and junctions) ahead. In Paramics network, stop lines are used to pass a vehicle from one link to another and a vehicle has to pass through an exit stop line and then an entry stop line. If the stop lines are correctly placed, and in line with each other, a vehicle does not have to change direction when changing links. Kerb points represent the inside and outside edge of the road surface. The position of the kerb points affects the default positions of the stop line points and the entry points of the link.

4.4.2 Reinforcement Learning

Reinforcement learning addresses the question of how an autonomous agent that senses and acts in its environment can learn to choose optimal actions to achieve its goals. Each time the agent performs an action in its environment, a trainer may provide a reward or penalty to indicate the desirability of the resulting state. For example, when training an agent to play a game the trainer might provide a positive reward when the game is won, a negative reward when it is lost, and zero reward in all other states. The task of the agent is to learn from this, indirect, delayed reward, to choose sequences of actions that produce the greatest delayed reward *(Mitchell, 1997)*.

4.4.2.1 Reinforcement-Learning Model

In the standard reinforcement-learning model, an agent is connected to its environment via perception and action, as depicted in **Figure 4-9**. On each step of interaction the agent receives as input, i, some indication of the current state, s, of the environment; the agent then chooses an action, a, to generate as output. The action changes the state of the environment, and the value of this state transition is communicated to the agent through a scalar reinforcement signal, r. The agent's behavior, B, should choose actions that tend to increase the long-run sum of values of the reinforcement signal. It can learn to do this over time by systematic trial and error, guided by a wide variety of algorithms that are the subject of later sections of this paper *(Kaelbling et al., 1996)*.

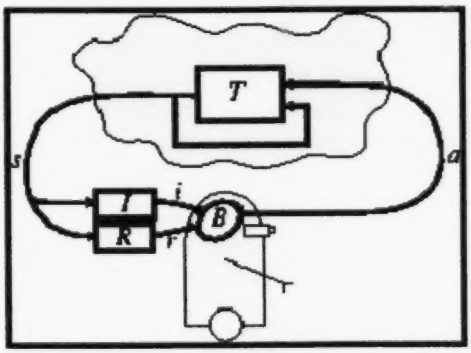


Figure 4-9 The Standard Reinforcement-Learning Model

Formally, the model consists of:

- A discrete set of environment states, $\mathcal{S}$;
- A discrete set of agent actions, $\mathcal{A}$; and
- A set of scalar reinforcement signals; typically {0, 1}, or the real numbers.

The figure also includes an input function I, which determines how the agent views the environment state; we will assume that it is the identity function, that is, the agent perceives the exact state of the environment.

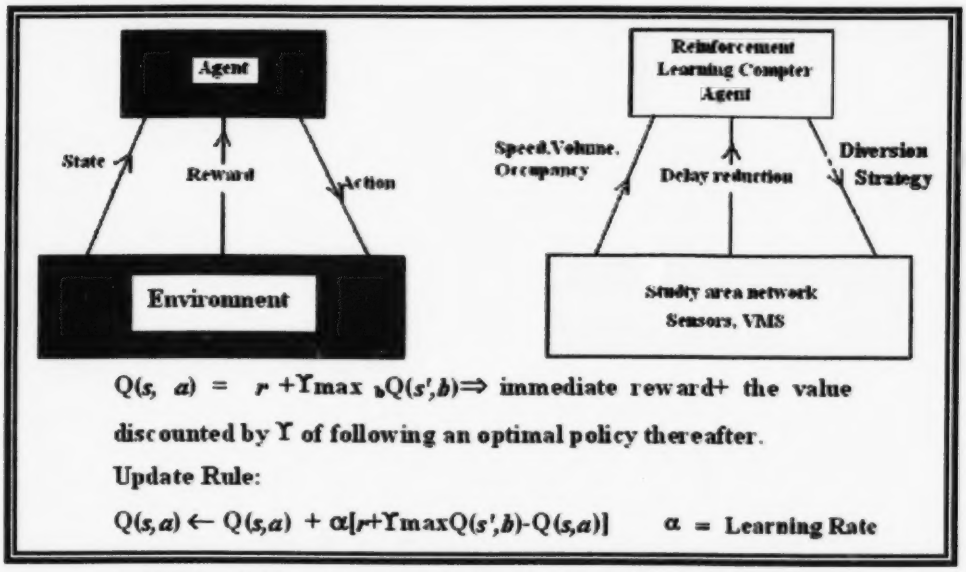


Figure 4-10 An Agent Behavior in its Environment

As shown in **Figure 4-10**, the agent takes sensory input from the environment and produces as output, actions that affect it. The interaction is an on going non-terminating one.

The agent exists in an environment described by some set of possible states S. It can perform an action A. Each time it performs an action $a_{t\,in}$ some state s_t the agent receives a real- valued reward rt that indicates the immediate value of this state action transition *(Mitchell, 1997)*. The agent that interacts with the traffic network is shown on the right side. Here, the agent receives information about the traffic such as speed, volume, and occupancy from the detectors and based on this information it selects some action of diverting a certain percentage of traffic by means of display on the VMS. As a result of the action, the agent receives a reward of delay reduction.

The task of the agent is to learn a control policy, $\pi : S \rightarrow A$, which maximizes the expected sum of these rewards, with future rewards discounted exponentially by their delay. This is expressed as:

$$V^{\pi}(s_t) = r_t + \gamma^1 r_{t+1} + \gamma^2 r_{t+2} \ldots \equiv \sum_{i=0}^{\infty} \gamma^i r_{t+i} \qquad (4\text{-}1)$$

Where $V^{\pi}(s_t)$ is the value of the state s_t , r_t is the immediate reward and Υ is the discount factor. We require that the agent learns the policy π that maximizes $V^{\pi}(s)$ for all states s. We call this policy an optimal policy and denote it by π^*.

$$\pi^* \equiv arg_{\pi} \max V^{\pi}(s), (\forall s) \tag{4-2}$$

It is difficult to learn the optimal policy function π^* for an arbitrary environment. The optimal action in state s is the action a that maximizes the sum of the immediate reward r(s,a) plus the value V* of the immediate successor state, discounted by Υ.

$$\pi^*(s) = \arg\max_a [r(s,a) + \gamma V^*(\delta(s,a))] \tag{4-3}$$

δ(s,a) denotes the state resulting from applying action a to state s. Thus, the agent can acquire the optimal policy by learning V*, provided it has perfect knowledge of the immediate reward function r and the state transition function δ. In many practical problems, such as traffic control, it is impossible for the agent or its human programmer to predict in advance the exact outcome of applying an arbitrary action to an arbitrary stage. In cases where either δ or r is unknown, the evaluation function Q defined below proves to be appropriate for the agent to use.

$$Q(s,a) = r(s,a) + \gamma V^*(\delta(s,a)) \tag{4-4}$$

The evaluation function is defined so that its value is the maximum discounted reward that can be achieved starting from state s and applying action a as the first action. In other words the value of Q is the reward received immediately upon executing action a from state s, plus the value (discounted by Υ) of following the optimal policy thereafter.

The Q(s,a) in **Equation (4-4)**, is exactly the quantity that is maximized in **Equation (4-3)** in order to choose the optimal action a in state s. Therefore the **Equation (4-3)** can be rewritten as:

$$\pi^{*(s)} = \arg\max_a Q(s,a) \tag{4-5}$$

Equation (4-5) shows that if the agent learns the Q function instead of the V* function, it will be able to select optimal actions even when it has no knowledge of the functions r and δ. It is clear from the above equation that the agent need only to consider each available action a in its current state s and choose the action that maximizes Q(s, a).

Q-Learning, which is initially introduced by Watkins, is a simple but efficient Reinforcement Learning algorithm that does not need a model of its environment and can be used on-line. Q-Learning algorithms work by estimating the values of state-action pairs. The value Q(s, a) is defined to be the expected discounted sum of future payoffs obtained by taking action a from state s and following an optimal policy thereafter. Once

these values have been learned, the optimal action from any state is the one with the highest Q-value *(Kaelbling, 1996)*.

Figure 4-11 shows an example to illustrate the concept of the Q-Learning.

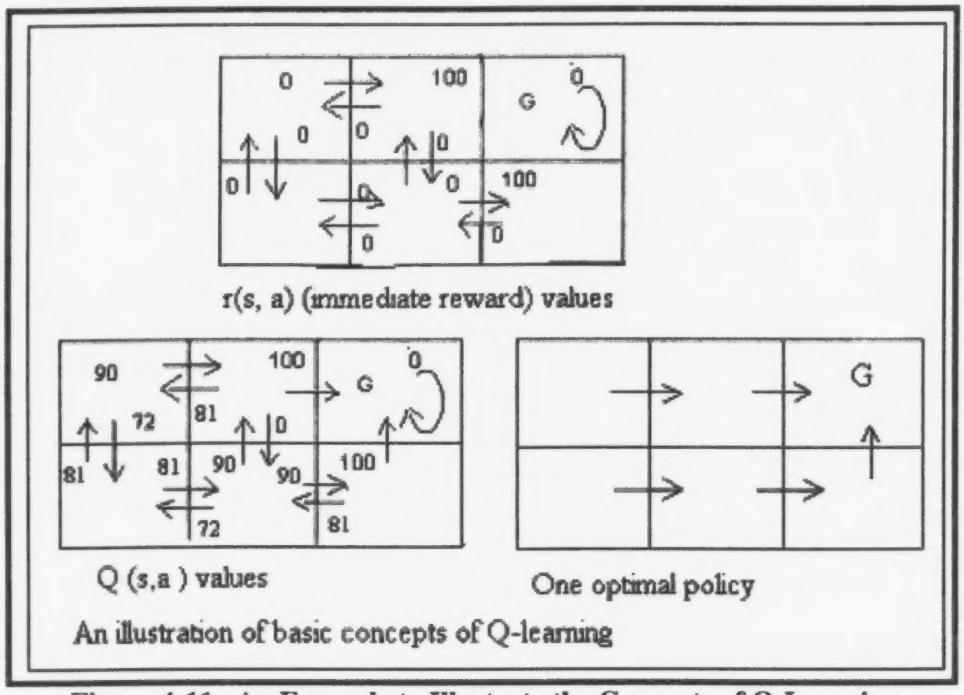


Figure 4-11 An Example to Illustrate the Concepts of Q-Learning

Each grid square represents a distinct state, each arrow a distinct action. The immediate reward function, r (s, a) gives reward 100 for actions entering the goal state G, and zero otherwise. Value of Q (s, a) follows from r (s, a), and the discount factor $\gamma = 0.9$. An optimal policy, corresponding to actions with maximal Q values is shown.

4.4.2.2 An Algorithm for Learning Q

Learning Q function corresponds to learning the optimal policy. The relationship between Q and V^{π} is,

$$V^{*}(s) = \max_{a'} Q(s, a') \tag{4-6}$$

This allows rewriting Equation (4-4) as:

$$Q(s, a) = r(s, a) + \gamma \max_{a'} Q(\delta(s, a), a') \tag{4-7}$$

This recursive definition of Q provides the basis for algorithms that iteratively approximate Q *(Watkins, 1989)*. For non-deterministic environment functions, the above training rule is modified as follows:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \tag{4-8}$$

Where α represents the learning rate [0 to1], which controls convergence, Υ is the discount factor and t is the time step. The discount factor makes rewards earned earlier more valuable than those received later and its value lies between 0 and 1.

After being initialized to arbitrary numbers, Q-values are estimated on the basis of experience as follows:

1. From the current state s, select an action a. This will cause a receipt of an immediate payoff r, and arrival at a next state s'.

2. Update Q(s,a) based upon this experience as follows:

3. Small changes in Q(s,a) = α[r+Υmax $_b$Q(s',b)-Q(s,a)]

4. Go to 1 if s is not terminal, with s $\leftarrow$ s'

This algorithm is guaranteed to converge to the correct Q-values with the probability one if the environment is stationary and depends on the current state and the action taken in it; called Markovian, where a lookup table is used to store the Q-values, every state-action pair continues to be visited, and the learning rate is decreased appropriately over time. This exploration strategy does not specify which action to select at each step. In practice, a method for action, called the Boltzmann distribution strategy, is usually chosen that will ensure sufficient exploration while still favoring actions with higher value estimates. This method learns the values of all actions, rather than just finding the optimal policy. This knowledge is expensive in terms of the amount of information, which has to be stored, but it does bring benefits. Q learning is exploration insensitive, any action can be carried out at any time and information is gained from this experience.

4.4.2.3 Exploration and Exploitation

ε - Greedy Method: One important decision in reinforcement learning is exploration vs. exploitation with respect to the policy to be adopted for selecting action in any state. One strategy would be to select the action a that maximizes the Q(s,a), thereby exploiting its current approximation Q' . We call this a greedy action and it will lead to exploitation. However with this strategy the agent runs the risk that it will over commit to actions that are found during early training to have high Q values. Instead, if a non-greedy action is taken, it is called exploration because this enables the agent to improve the estimate of the non-greedy action's value. A simple action selection method is the ε- greedy action in

which most of the time a greedy action will be taken and a non-greedy action with a small probability ε *(Millar, 1999).*

Softmax Action Selection: Although ε- greedy action selection is an effective and popular means of balancing exploration and exploitation in reinforcement learning, one drawback is that when it explores it chooses equally among all the actions. This means that it is as likely to choose the worst appearing action, as it is to choose the next-to-best action. In tasks where the worst action is very bad, this may be unsatisfactory. The solution is to vary the action probabilities as a graded function of estimated value. The greedy action is still given the highest selection priority, but all others are ranked and weighted according to their value estimates. These are called Softmax action selection rules. The most common method uses a Gibbs, or Boltzmann, distribution *(Sutton and Barto, 1998).*

$$P(a_i / s) = e^{Q(s,a_i)/\tau} / \sum_j e^{Q(s,a_j)/\tau} \qquad (4\text{-}9)$$

Where P(ai / s) is the probability of selecting action a_i, given that the agent is in state s, τ is a positive parameter called the *temperature* . High temperatures cause the actions to be all nearly equiprobable. Low temperatures cause a greater difference in selection probability for actions that differ in their value estimates. In the limit as $\tau \rightarrow 0$, Softmax action selection becomes the same as greedy action selection. Setting τ requires knowledge of the likely action values and of powers of e.

This project adopts a simple ε- greedy exploration method with a gradually decreasing rate of exploration so that at the beginning mostly exploration when the agent doesn't know much and towards the end of the learning more of exploitation takes place. Gradual shifting is required to ensure that the entire state space is covered during learning. One possibility for selecting an action; pick the "best" action with probability $P = 1 - e^{-E a}$, suggested by Michael Gasser *(2004)*. Where a is the number of training samples (the "age" of the agent). **Figure 4-12** and **Figure 4-13** show how the probability of selecting the "best" action depends on age when E is 0.1 and 0.01.

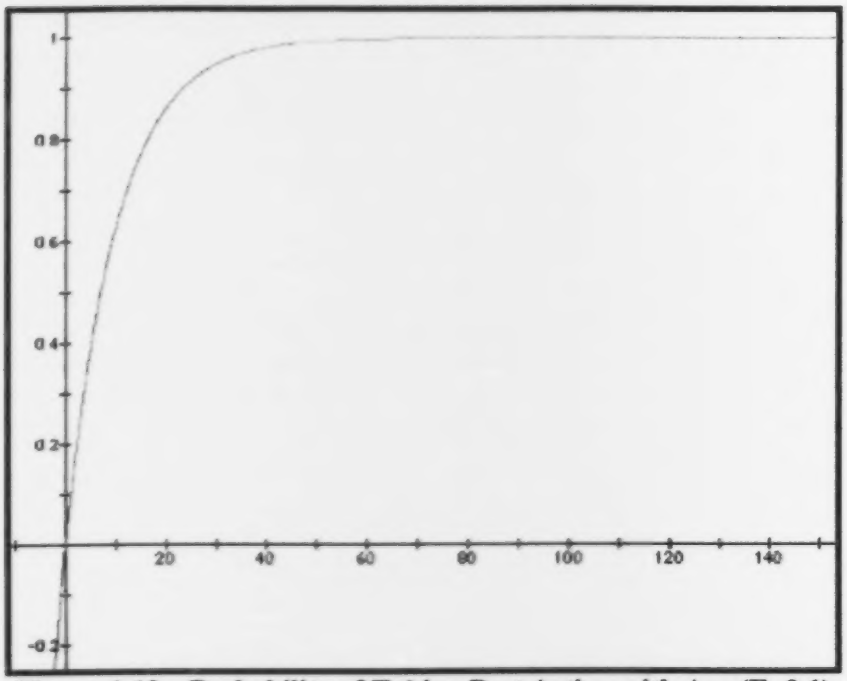


Figure 4-12 Probability of Taking Best Action with Age (E=0.1)

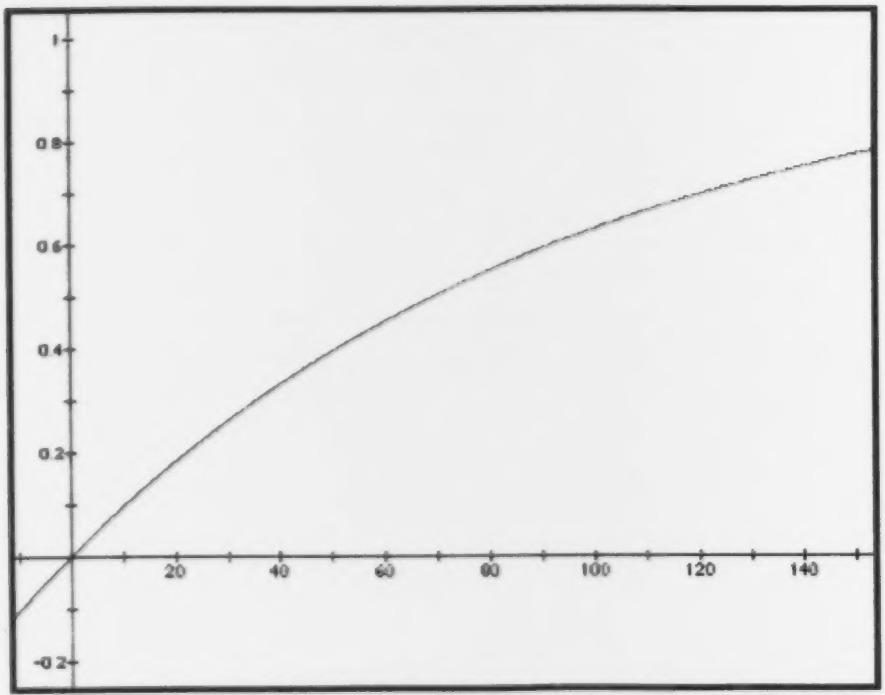


Figure 4-13 Probability of Taking Best Action with Age (E=0.01)

For the state space defined in this project a value of E = 0.001 has been found to be appropriate. For a very large state space, a still smaller value say E =0.0005 would be suitable.

4.4.2.4 Learning Rate

The learning rate controls how fast we modify our estimates. One expects to start with a high learning rate, which allows fast changes, and lowers the learning rate as time progresses. During Learning, the various states will be visited in an asynchronous manner. In order to keep the error due to this asynchronism negligible, one must make sure that the step size (learning rate) is small. The step size rule suggested by Darken et al (1992) given below has been selected for this project.

$$\alpha^{n+1} = T_1 /[1 + n^2 /(1 + T_2)] \tag{4-10}$$

Where α^{n+1} is the learning rate for the n^{th} visit to a particular state. T_1 is the starting value for α and T_2 could be set as desired to control the decay of α. Possible values of T_1 and T_2 are 0.1 and 10^4 respectively.

4.4.2.5 Generalization from Training Examples

To estimate the Q value for unseen state action pairs it requires generalization from those states that have already been visited. Function approximation techniques like Back-Propagation algorithm or Cerebellar Model Articulation Controller (CMAC) *(Smart, 2002 and Sutton, 1996)* could be used for the generalization. CMAC algorithm has proven to be popular for value-function approximation work. The algorithm splits the input space into number of overlapping receptive fields, each with an associated response. Any given input will lie within a small number of these fields, known as activated fields. The response of the CMAC to this input is simply the average response of the activated cells. **Figure 4-14** illustrates this with simple CMAC with a two-dimensional input space and 4 receptive fields or tiles. An input state indicated by "•" will activate one cell each carrying a weight, from each of the tiling. The response of the CMAC neural network to a given input is the average of the responses of the receptive fields exited by that input and is not affected by the other receptive fields. In the case Q-Learning using CMAC, training is performed based on the Q Learning update rule described earlier as follows to update the weight vector W.

$$Q(s,a)_{new} = Q(s,a)_{old} + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a)_{old}) \tag{4-11}$$

$$W_{new} = W_{old} + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a)_{old}) \tag{4-12}$$

Generalization means CMAC gives similar outputs when given similar inputs. On the other hand it should be capable of producing independent outputs for dissimilar inputs.

Generalization is important because learning cannot cover every state-action. CMAC achieves robustness through generalization. A CMAC designer can control the amount generalization by controlling the number of tiles. If only one tile, then there will not be any generalization and a large number of tiles means much generalization.

This project employs a CMAC receptive field organization as suggested by the University of Hampshire in their implementation of CMAC *(Glanz et al., 1991, Miller and Ghanz, 1996, Albus, 1975)* suited for software implementation.

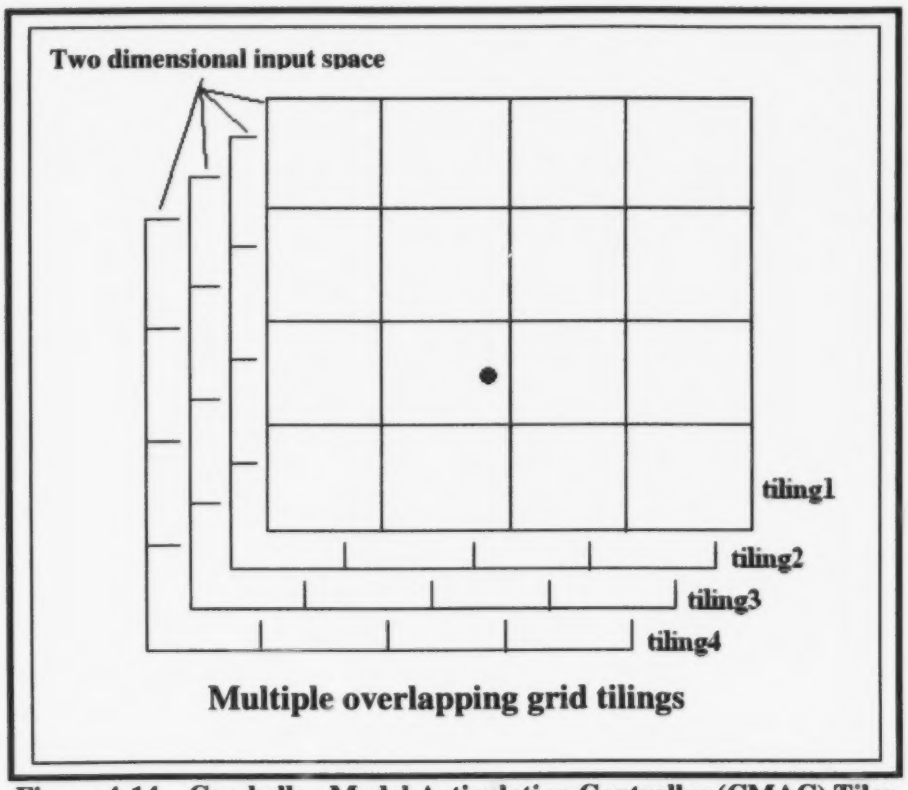


Figure 4-14 Cerebellar Model Articulation Controller (CMAC) Tiles

The main features of this implementation are given below:

- The total collection of receptive fields is divided into C tiles or layers each of which represent a parallel N dimensional hyperspace for a network with N inputs.
- The receptive fields in each of the layers have rectangular boundaries and are organized so as to span the input space without overlap.
- Any input vector excites one receptive field from each layer.
- Each of the layers of receptive fields is identical in organization, but each layer is offset relative to the others in the input hyperspace.

- The width of the receptive fields produce input generalization, while the offset of the adjacent layers of receptive fields produces input quantization.

- The ratio of the input generalization to the input quantization must be equal to C for all dimensions of the input space. This will ensure that only a fixed number C, of receptive fields is excited by any input space

In case of systems with large state space, the total number of receptive fields required to span the entire state space would be huge for many practical problems. On the other hand only a small percentage of this state space will be visited in reality. Thus a look up table, which stores unique information for each receptive field, is not necessary. Following this logic, a form of pseudo-random hashing will be included to store only information about receptive fields that have been excited during previous training.

The adopted CMAC computation is shown diagrammatically in **Figure 4-15**, which uses the following steps:

- N dimensional real valued input vector $S = < S_1, S_2, S_3, S_4, \ldots, S_N, >$, received from the sensors be converted to normalized integer input vector $S' = < s_1, s_2, s_3, s_4, \ldots, s_N, >$ by dividing each component S of the input vector by an appropriate quantization parameter.

- The receptive fields are arranged in a geometrically regular way, such that the coordinates of the C excited receptive fields for any input can easily be determined in software or generated in hardware, without having to compare to independent coordinates stored for each of the very many receptive fields. This is achieved by keeping the width of receptive field equal to C and the receptive field centers are arranged in a lattice defined by displacement vector $D =< d_1, d_2, d_3, d_4, \ldots, d_N, >$ such that the coordinate of the i^{th} receptive field in the reference hypercube was $< (i^* d_1) \% C, (i^* d_2) \% C, \ldots, (i^* d_N) \% C \qquad i=1,2\ldots,C$

  Where % is the modulus operator.

- Obtain a vector address A' of the C receptive fields which contain the input point S'.

$$A_i' = < s_1 - ((s_1 - i^* d_1) \% C, s_2 - ((s_2 - i^* d_2) \% C, \ldots s_N - ((s_N - i^* d_N) \% C >$$
$$= < a_{i1}, a_{i2}, \ldots, a_{iN}, > \quad i = 1,2,\ldots,C$$

- Since the total number of receptive fields in a state space of dimension N can be quite large, the next step is to obtain scalar physical address A_i of the actual adjustable weights to be used in the output computation from the virtual address A_i'.

  $$A_i = h (a_{i1}, a_{i2}, \ldots, a_{iN})$$

h(...) represents any pseudo random hashing function which operates on the components a of the addresses of the receptive fields, producing uniformly distributed scalar addresses in the physical weight memory of size M.

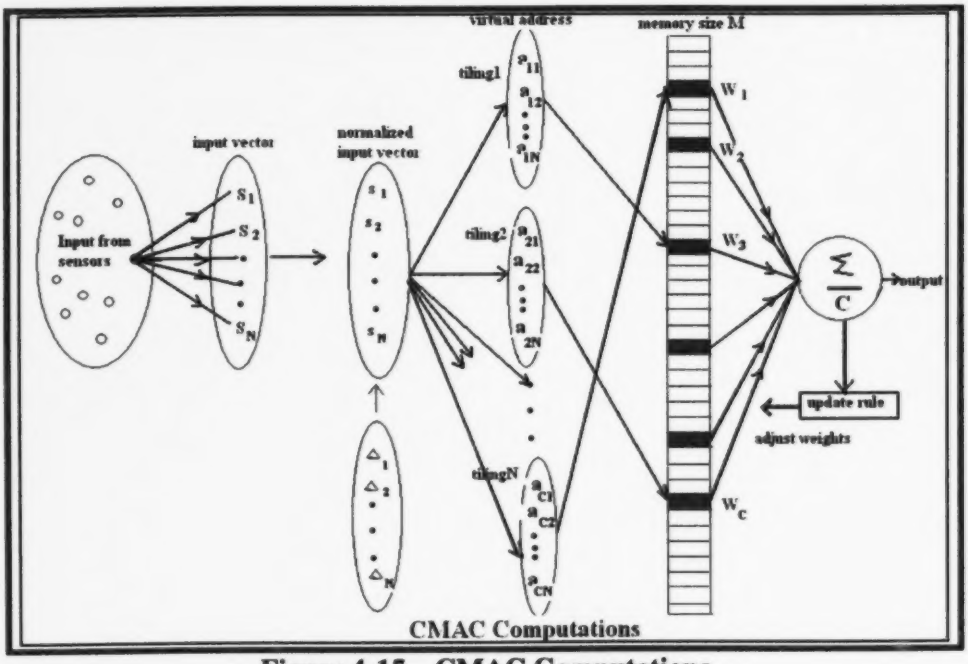


Figure 4-15 CMAC Computations

5 TRAVEL DEMAND DATA

The volume data for the study were obtained from two sources. The first data set consisted O-D volumes was obtained from the Transportation Tomorrow Survey (TTS), one-day travel survey conducted in the Greater Toronto and surrounding areas. Travel demand in Paramics is defined by a matrix of origin to destination trips specified in the demands file. Trips are released from and arrive at geographic areas called zones. The zones specified for the study area network in Paramics is shown in **Figure 5-1**. Zones were made at all entry and exit points (separate zones to represent each on-ramps and off-ramps of Collector and Express) of the study area so that the area is delineated from the rest of the region. The 2001 Transportation Tomorrow Survey (TTS), one-day travel survey conducted in the Greater Toronto and surrounding areas, was used to determine zonal flows and vehicle proportions.

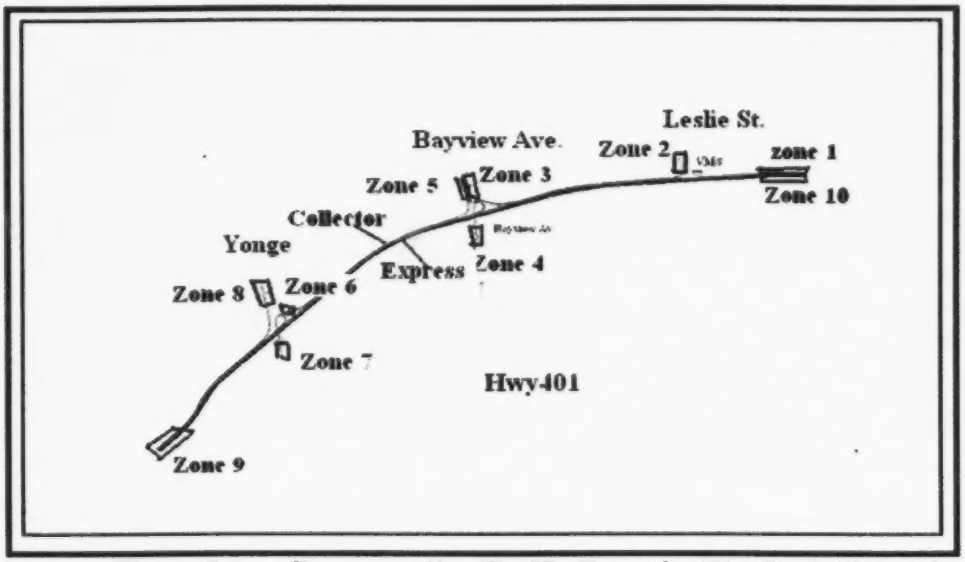


Figure 5-1 Corresponding Traffic Zones for The Study Network

A static assignment based transportation modeling software- EMME/2 *(1998)*, was employed to develop the demand matrix used to run the micro-simulator. A full-scale assignment for the morning peak for the whole Greater Toronto Area was conducted. An advanced EMME/2 feature, 'traversal operator', which computes trip interchanges into and out of any sub-area of a sub-network, was used to develop an O-D matrix for the sub-network consisting of the subject Express/ Collector area. A cordon around the subject freeways was defined and traffic crossing the gateways constituted O-D flows for the sub network.

The traversal matrix obtained for these zones from EMME/2 is shown in **Table 5-1**. Based on available information the zones 1, 2 and 10 as defined in the study area network in Figure 2 are combined into a single zone 1 while defining zones in EMME/2 for the generation of traversal matrix. Because of this fact, the O-D data obtained from EMME/2

cannot be directly used to derive the O-D matrix required for our purpose. However this table would be useful for gaining volume distribution between zones.

Table 5-1 Traversal Matrix Generated from EMME/2 Using 2001 TTS data

Zone of Origin \ Zone of Destination	1, 2, & 10	3	4	5	6	7	8	9	Total
1, 2 & 10	0	1960	0	0	1395	0	0	9450	12805
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	18	0	0	210	228
5	0	0	0	0	246	0	0	885	1131
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	400	400
8	0	0	0	0	0	0	0	1393	1393
9	0	0	0	0	0	0	0	0	0
Total	**0**	**1960**	**0**	**0**	**1659**	**0**	**0**	**12338**	**15957**

The second data set consisted of live data obtained from the detectors at the study area. **Table 5-2** shows the average of several morning peaks one-hour volumes collected from the detectors in the study area. This aggregated information is gathered from the ICAT (ITS Centre and Testbed) database and platform maintained at the University Toronto using the live detector and video data transmitted from MTO and City of Toronto.

Table 5-2 Peak Period Detector Volumes from the Study Area

Detector	Location	7-Sept.	8-Sept.	9-Sept.	10-Sept.	13-Sept.	14-Sept.	15-Sept.	Avg.
401DE0470DWR	S Leslie - 401 W RN	708	704	686	747	699	778	752	725
401DE0180DWC	West of Leslie CL	6412	7366	6002	7288	6486	7340	6661	6857
401DE0180DWE[1]	West of Leslie EL	---	---	5184	5711*	5054	5386	5344	5242
401DE0170DWC	East of Bayview CL	---	4631	4856	4528	4144	4326	4691	4529
401DE0170DWE	East of Bayview EL	---	---	4522	4732	5321	5332	4657	4913
401DE0060DWT	West of Leslie TL	---	---	3181	3692	3725	3927	3558	3617
401DE0160DWC	At Bayview Ave. CL	4512	4434	3970	3098	2819	2802	3038	3525
401DE0160DWE	At Bayview Ave. EL	4994	5138	6763	6924	7181	7345	7080	7059
401DE0450DNR	W 401 - Bayview N RX	1243	1293	982*	1841*	1039	1336	1043	1191
401DE0440DWR	N Bayview - 401 w RN	186	204	185	192	209	222	193	199
401DE0420DWR[2]	S Bayview - 401 W RN	---	---	---	---	---	492	512	502
401DE0130DWC	At Yonge CL	2033	2166	1737	2201	2085	2019	1998	2034
401DE0130DWE	At Yonge EL	3249	3276	2887	4574	4384	4716	4631	3960
401DE0400DNR[3]	W 401 - Yonge St. N RX	---	---	---	1589	1443	1497	1745	1562
401DE0120DWC[4]	W of Yonge CL	---	---	---	---	---	---	---	---
401DE0120DWE[4]	W of Yonge CL	---	---	---	---	---	---	---	---

[1] *Data from July 2004*
[2] *It was not working during study periods, data from July 2004*
[3] *No data available, interpolated volume*
[4] *No data available*
* *Extreme values neglected*

As discussed earlier we cannot use volumes generated from EMME/2 directly. Similarly the detector volumes in **Table 5-2** give only the volumes crossing that particular detector location which does not correspond to the O-D volumes. Therefore some approximation and assumptions were made to obtain the O-D volumes to use in this study. Using the currently available detector volumes from Table 2 and based on the O-D volumes obtained from EMME/2 shown in **Table 5-1**, an O-D matrix as shown in **Table 5-3** was formed. Because of the unavailability of some detector data, it may not exactly match with the actual value. Volume from the detector 401DE0180DWC is 6857 and includes demand from zone 1 (on ramp at Leslie) and zone 2 (East of Leslie). We have data from

the Leslie on-ramp as 725, which is assumed as through traffic hence this value is shown as volume from zone 2 to zone 9. Remaining volume of 6132 is assumed to be the total volume that is released from zone 1. As there is no transfer possible from Express to Collector between Yonge and Leslie, the volume from detector 401DE0180DWE, which is equal to 5242, can be taken as the demand from zone 10 to zone 9. The sum of the demands from zones 1,2 and 10, which is equal to 12099 (6132+725+5242) is matching with that of the demand obtained from EMME/2, which is 12805. But the off-ramp volumes at Yonge Street and Bayview Avenue have lower numbers as compared to the EMME/2 volumes (1191 and 1562 as against 1960 and 1395). This is expected since people will try to avoid using these ramps due to ongoing construction in that area. The reliability of the detectors at the work zone area is also in question. However for the simulation, the data obtained from the detectors were used instead of that from EMME/2. Therefore volume from zone 1 to zone 3 is 1191 and that from zone 1 to zone 6 is 1562. The remaining volume of 3379 from zone 1 will be the through traffic to zone 9. Demands from zones 4 and 5 (on ramps at Bayview Avenue) to zone 9 are taken same as obtained from the corresponding detectors as 199 and 502 respectively. Also demands from zones 7 and 8 were taken from the traversal matrix (as there is no detector data available) with slight reduction to match with the reduced on ramp demand due to construction.

Table 5-3 O-D Matrix for the simulation

		Zone of Destination										
	From/To	1	2	3	4	5	6	7	8	9	10	Total
Zone of Origin	1	0	0	1191	0	0	1562	0	0	3379	0	6132
	2	0	0	0	0	0	0	0	0	725	0	725
	3	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	199	0	199
	5	0	0	0	0	0	0	0	0	502	0	502
	6	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	300	0	300
	8	0	0	0	0	0	0	0	0	1200	0	1200
	9	0	0	0	0	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0	0	5242	0	5242
	Total	**0**	**0**	**1191**	**0**	**0**	**1562**	**0**	**0**	**11547**	**0**	**14300**

6 IMPLEMENTATION OF DEVELOPED SYSTEM IN SIMULATION ENVIRONMENT

The Paramics network of the study area is provided with the above demand and the features like cost factor, signposting distance, stop lines, and kerb points etc. are adjusted with several trial runs to simulate the field conditions and also to ensure smooth flow of traffic as much as possible. Each simulation is run for one and half hours with a warm up period of half an hour. During the warm up period the traffic will be gradually released with a five-minute profile reaching the full release rate at end of the warm up period. Thereafter, the rate of release is kept constant for the actual simulation time of one hour. The Q Learning agent is activated after the warm up period and at every 5-minute interval it receives the state of the network as described earlier and selects an action based on the policy adopted and implements this action for the next 5 minutes. At the end of this period the learning agent will receive the reward as explained earlier and it updates the Q-value of that state action pair visited. This is continued till the end of the simulation and is called first iteration. The probability of exploration will be reduced gradually from one iteration to the next. Similarly the learning rate or step size also will be reduced gradually as the number of visit to the state action pair. The updated Q-value of each of the state-action pair visited and the number of visits to that pair is stored in separate files. Another file stores the state visited, action performed, number of visits so far, reward from that state and the updated Q-value. The learning should be continued till the learning rate for all state action pair approaches zero and exploration ceases to exist. In other words the Q-value for all the state-action pair reaches its maximum value. The test runs were done after every 100 iterations and the Paramics outputs were taken for the overall evaluation as discussed earlier. During the test runs, the agent will take all greedy actions at every state to utilize the knowledge obtained so far *(Jacob and Abdulhai, 2004).*

6.1 VMS Control

In order to control the flow of traffic through the work zone area appropriate messages has to be displayed. The function of the learning agent in this project is to provide this information at a regular interval say 5-minutes based on the state of the traffic flow on both Collector and Express. The agent learns this optimal strategy using the simulator. In the Paramics network the VMS and the detectors are coded at the location corresponding to the site locations. Unlike detectors the functionality of the VMS cannot be achieved without providing the logic required to direct each vehicle as per the requirement. The entire Q- Learning program with all the logic to obtain states, implementing actions based on the VMS display, obtaining the rewards, updating Q-values, output operations etc are all coded as a plugin program and will be accessed by the Paramics during the simulation through its API.

Paramics API has got its own basic functions and the user could implement the required functionality by providing the control logic either as override functions or as overload

control functions. It also has Call back functions that will provide information about some of the attributes of the current vehicle and its environment, such as the link it is traversing, as well as providing simple debug facilities.

In this present set up the VMS will be provided with a message appropriate to divert 0, 20, 40, 60, 80 or 100 percent of the through traffic from the Collector to the Express. During learning if an action of diverting a particular percentage is selected the program is provided with the logic to exactly divert the correct percentage and hence the effect of that action (reward) will be obtained after implementing action. In order to facilitate this, the plugin is provided with the following information:

- The number of VMS on the network- vms_ control_points.
- The link on which a particular VMS is located- vms_control_links.
- Destination zones for the diverted traffic- capture_zones.
- The links along the collector and express to guide the traffic along its path.

In the study network, VMS is provided at Leslie upstream of the Collector to Express transfer. At every time step (half a second) each vehicle in the network will be checked for its current link and if it is on the VMS link and if its destination is on the capture zone then this vehicle will be considered for route choice check. In the route choice decision function the vehicle will be either flagged for Collector or Express based on the message display on the VMS and also based on the number already considered during the time period (5min.). Decision will be taken to effectively divert the specified percentage. Depending the flag of the vehicle it will be directed along Collector or Express by specifying the *nextlink* at each diversion point for that vehicle.

However in reality, the reaction of a motorist to a VMS display depends on many factors like driver awareness, aggression, familiarity, trust on the VMS, variation in their perceived cost etc. The task of the manager is to display a message appropriate for the required diversion. A Reinforcement Learning agent set for continuous on-line learning, if the diversion percent is not as expected, the reward obtained will be less and that will automatically adjust the Q value of that state action pair and as a result next time the agent will try some better action. Hence, agent will be will be able to learn about the uncertainties through on-line learning.

6.2 States for the Q-Learning

The state of the network at any instant speed, volume, and occupancy could be obtained from the loop detectors provided along the road. Paramics also has the ability to add loop detectors on any specified link and also to output the above values for any instant of time or whenever a vehicle crosses a detector. In this project, 3 variables are used to represent the current state of the network. They are:

- Average speed on the Collector;
- Average speed on the Express; and
- Current diversion percentage.

The average speed can be obtained by averaging the speeds of the vehicles crossing the detectors on the downstream of the diversion point. Current diversion percentage will be the action selected during previous time interval. In the case of Q-Learning, which employs a look up table to store the Q-values, the speed state is defined to fall into any of the bin. In order to keep the state space manageable 3 km/hr has been taken as the size of each bin, i.e. 0-<3 km/hr in state 1, 3-<6 km/hr as state1 and so on. In the case of CMAC implementation this number is considered as the quantization parameter for the state.

6.3 Actions for the Q-Learning

Action, here, is the percentage diversion to be applied. Six actions were considered. Action 0 for 0% diversion, 1 for 20%, 2 for 40% and so on with action 5 representing 100% diversion. This diversion percentage will be implemented during the next time interval and there after a new action will be selected. Depending upon the exploration rate a random action or the best action so far will be taken at each stage.

6.4 Reward for the Q-Learning

The reward can be either positive (+ve) or negative (–ve) based on whether a benefit (e.g. delay reduction) or penalty (e.g. delay) is taken as the outcome of the action. A reward function suggested by Papageorgiou et.al *(2003)* is utilized to represent the reward function. The total time spent by all the vehicles in the traffic network will be longer if, for any reason (e.g. due to lack of control measures or due to incident), the exit flows are temporarily lower, i.e. if vehicles are delayed within the network on their way to the destinations. As a consequence, any control measure or control strategy that can manage to increase the early exit flows of the network, will lead to a corresponding decrease of the total time spent. Consider a discrete time representation of traffic variables with discrete time index k= 0, 1, 2, 3… and time interval T. The aim is to apply control measures so as to minimize the total time spent Ts in the network over a time horizon K, i.e.

$$T_s = T\sum_{k=1}^{K} N(k) \qquad (6\text{-}1)$$

Where N(k) is the total number of vehicles in the network at time k.

A *traffic volume or flow q(k)* (in veh/hr) is defined as the number of vehicles crossing a corresponding location during the time period [kT, (k+1)T], divided by T. Traffic density ρ(k) (in veh/km) is the number of vehicles included in a road segment of length Δ. Mean speed v(k) (in km/hr) is the average speed at time kT of vehicles in a road segment.

$$N(k) = N(k-1) + T[d(k) - s(k)] \tag{6-2}$$

$$N(k) = N(0) + T\sum_{k=1}^{k-1}\{d(k) - s(k)\} \tag{6-3}$$

Substituting (6.3) in (6.1) we obtain

$$T_s = T\sum_{k=1}^{K}\left\{N(0) + T\sum_{k=0}^{k-1} d(k) - T\sum_{k=0}^{k-1} s(k)\right\} \tag{6-4}$$

The first two sums in the outer sum **Equation (6.4)** are independent of the control measures taken in the network; hence, minimization of Ts is equivalent to maximization of the following quantity:

$$S = T^2\sum_{k=0}^{K}\sum_{k=0}^{k-1} s(k) = T^2\sum_{k=0}^{K-1}(K-k)s(k) \tag{6-5}$$

Thus minimization of the total time spent in a traffic network is equivalent to maximization of the time-weighted exit flows. In other words, the earlier the vehicles are able to exit the network (by appropriate use of the available control measures) the less time they will have spent in the network.

In this project, a time-weighted exit flow during the period is taken as the reward. The period of 5 minutes is divided into 5 steps of 1 minute each and the outflows are gathered. Here: K= 5, T = 1 min. and Reward = 5*s(0) + 4*s(1) +3* s(2) +2* s(3) + 1* s(4). The Q-Learning agent tries to maximize this reward at each state of the network.

6.5 Evaluation Measures

In order to evaluate the feasibility of the approach, the result of using the learning agent will be compared with the result obtained from the base scenario- the default actions. The testing was done after a sufficient number of iterations with various state values. For this purpose the Paramics simulation will be run for the same network twice- once with the learning agent active and then without using the learning agent. In the test runs, the action selected will be the one with maximum Q value. The network for the testing phase is provided with the measurement file to gather economic info to output the summary statistics from modeler that give network wide information about delay, queuing, number

of vehicles on the network, distance traveled and mean speed. The column headers in the output file are listed and described as follows:

```
## ------------------------------------------------------------------------
##         |   current     |      all vehicles                 |  bus-specific
## time    |   nv  mean    | travl total total  mean stop|
##         |       speed   |  time  dist    nv speed time| stops time  on off
## ------------------------------------------------------------------------
```

- *Time:* each line has a timestamp. The file will be refreshed every minute of simulation, regardless of any other settings;
- *Current:* relating to vehicles that are being simulated at the instant of the timestamp;
- *nv:* the current number of vehicles;
- *Mean speed:* the average speed in either miles/h (US and UK mode) or in km/h (metric mode);
- *All vehicles:* relating to all vehicles (all trips) released since the simulation began, including those that have not yet reached their destination (the current vehicles);
- *Travel time:* mean travel time (seconds) per vehicle in the simulation, up to the current timestamp. Here travel time means the time taken to get from origin to destination (or from origin to where the vehicle is at the present timestamp), zero time would require infinite speed;
- *Total destination:* total cumulative distance traveled up to the current timestamp in the simulation, either in meters (UK and metric mode) or in feet (US mode);
- *Total nv:* the total cumulative number of vehicles in the simulation up to the current timestamp;
- *Mean speed:* the average speed for all vehicles up to the current timestamp in the simulation, either in miles/hr (US and UK mode) or in km/hr (metric mode);
- *Stop time:* mean stop time per vehicle in seconds, where "stopped" is defined as a vehicle traveling at less than the queue speed which is defined in the Configuration Manager. For example, a queue speed of 5m/s (~ 11mph) means that vehicles moving faster will not be counted as being queued; and
- *Bus-specific:* values relating to ALL PT vehicles released since the simulation began.

All the above parameters could be used to compare the effectiveness of the model. In this project, the parameters, which relate to the all vehicles viz. the average travel time, total distance, mean speed and stop time etc. were used for the evaluation.

- % Reduction in average travel time = 100* (Average travel time with base case - Average travel time with Q-Learning agent) / Average travel time with base case;
- % Increase in distance = 100* (total distance with base case - total distance with Q -Learning agent) / total distance with base case;
- % Increase in mean speed = 100* (mean speed with base case – mean speed with Q-Learning agent) / mean speed with base case; and
- % Reduction in stop time = 100 * (stop time with base case - stop time with Q-Learning agent).

6.6 Implementation in Simulated Environment

Two different agents- one using a simple look-up table to store the Q-values and the other with the elaborate CMAC implementation for function approximation were developed. Both these agents were provided with the learning scenarios and the outputs were obtained. Following are the parameters selected in this project based on the descriptions earlier.

Exploration: E =0.001
Learning rate: $T_1 = 0.1, T_2 = 104$
Discount Rate: $\Upsilon = 0.3$
CMAC:
Number of tiles or layers C = 16 (it is recommended that C to be the smallest power of 2 which is equal to greater than 4*N. Where N is the input dimension.)
Number of CMAC's = 6 (one for each action)

These two implementations are described next.

6.6.1 Look-up table implementation

This is suitable when the state space of the agent is small and the agent can be provided with all possible states during learning operation. It is simple to develop but the disadvantage is that for a large state space the memory requirement will be very high and if the agent comes across with a state, which it has not visited, the action will be inferior.

6.6.2 CMAC Implementation

In this setup a function approximation technique CMAC is incorporated in the original Q-Learning algorithm. Here instead of a lookup table the CMAC weights will be stored for each state action pair visited. This implementation is particularly advantages with large state space thereby eliminating the curse of dimensionality. Due to approximation, the result with states, which were visited during learning, may be slightly lower than that of the look up table approach but the performance with unvisited states will give out much better result than that of look up table approach.

6.6.3 Results

The data from the detectors at the study area indicated a diversion of 80% of the through traffic from Collector to Express at the transfer at west of Leslie. Hence the simulation with 80% diversion through out is considered for the base case. In order to compare fixed diversion rate against real time dynamic diversion, runs are also performed with various fixed diversion rates such as 0%, 20%, 40%, 60%, 80% and 100%. All these results are then compared with the result obtained using the Learning Agent. **Table 6-1** provides these results. Figures in brackets give the percent improvement by using the learning Agent as compared to the base case of 80% diversion.

Even though the CMAC simulation result is slightly inferior to that of the look up table approach, both Learning Agents have shown considerable improvements when compared with the base case and all other fixed diversion case.

Table 6-1 Results of Simulation with Q-Learning agents and various Fixed Diversion Percentages

	Percentage Diversion							
	0%	20%	40%	60%	80% (Base Case)	100%	Q-Learning Look UP Table	Q-Learning CMAC
Mean Travel Time (sec)	736.8	624.1	498.7	528.9	670	735.9	437.1 *(34.7%)*	466.9 *(30.3%)*
Total Vehicle Distance (km)	84051.9	90674.9	97875.6	96488	86489.2	82466.1	102953.8 *(19.0%)*	100584.97 *(16.3%)*
Mean Speed (km/hr)	26.5	32.4	41.0	37.8	27.1	24.6	47.4 *(74.9%)*	43.9 *(62.0%)*
Stop Time (sec)	6886	5341	3628	3901	5802	6727	2375 *(59.0%)*	2967.8 *(48.8%)*

Figure 6-1 shows the comparison of Mean Travel Time for the above scenarios and **Figure 6-2** shows that for the Mean Speed.

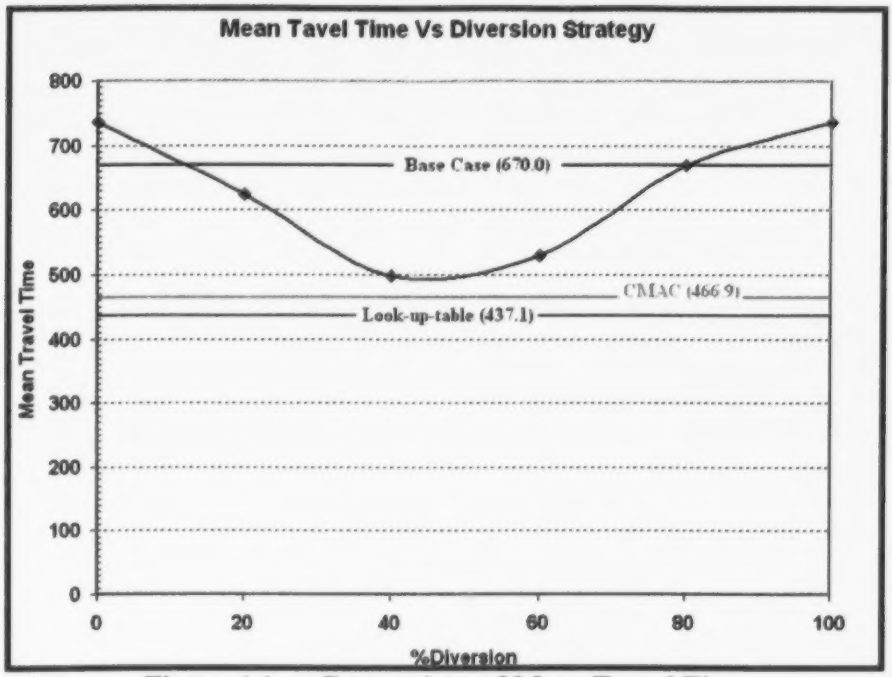


Figure 6-1 Comparison of Mean Travel Time

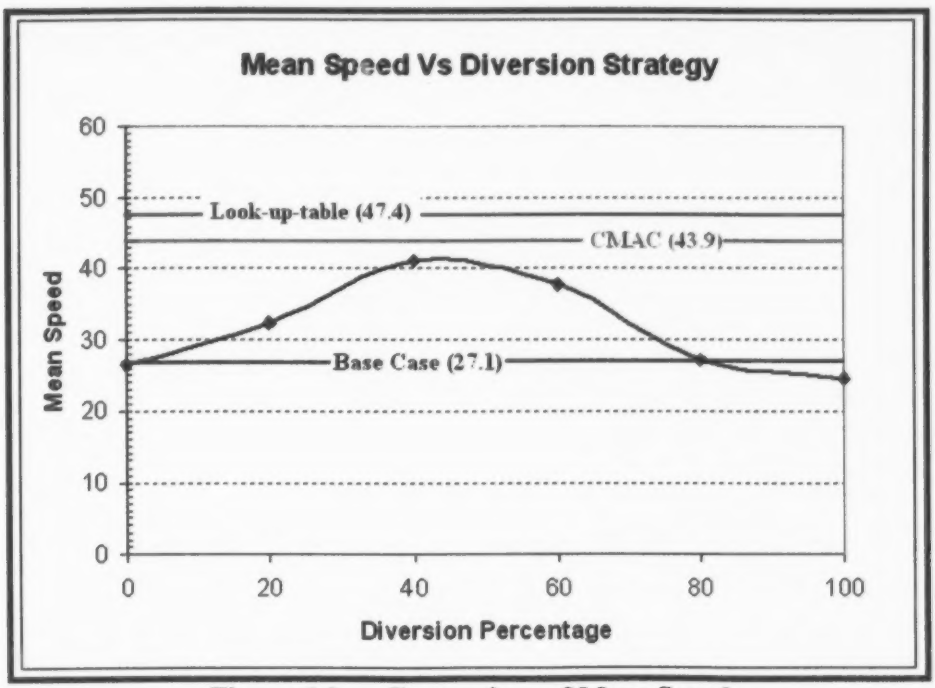


Figure 6-2 Comparison of Mean Speed

In both cases, it can be seen that the Q-Learning agent has outperformed all other fixed diversion scenarios. The order of performance among the fixed diversion scenarios for this test case is 40% performing best, followed by 60%, 20% and 80%.

Table 6-2 shows the results of the simulation assuming that drivers have no idea whatsoever about the work zone. In this case the improvement is very marginal. It gives light into the fact that, controlling traffic just at one point alone is not sufficient to achieve effective result for major congestions. In other words, if the drivers were informed of the situation well upstream of the construction zone and then if real time control is applied just at the foot of the congested area it will be more manageable.

Table 6-2 Results for Drivers without Prior Information about Work Zone

	40%	60%	80%	Natural	QL
Mean Travel Time (sec)	1412.4	1328.1	1372	1371	1314
Mean Speed (km/hr)	11.7	12.1	11.3	12	12.5

6.7 Analysis and Results Based on Occupancy States

During the October 26, 2004 meeting with MTO's staff, they raised certain issues regarding the accuracy of the data from the work zone area. Most of the detectors at the study area were substituted by the RTMS. It was highlighted that the volume data available from the RTMS is inferior to that of the occupancy data as regards to its accuracy especially during heavy traffic flow. Since the accuracy of the data obtained is an important factor for the developed methodology, it was decided to use occupancy data instead of volume to represent the state of the system.

There is a significant difference between the Paramics loop detector occupancy data and the real world loop occupancy data. In Paramics, the point data is gathered at every time step (normally 0.5 sec.) when an individual vehicle passes over the loop. This gives the time occupancy i.e. the time that particular vehicle is on the detector. This depends on vehicle length, detector length and vehicle speed. *Incomplete Occupancy* in Paramics is the time the vehicle is on the detector during a specific instant of time and Complete Occupancy defines the occupancy associated with the last vehicle that crosses a loop detector whereas in the real world, loop detectors provide the percentage occupancy. It is defined as the percentage of time the loop is occupied by the vehicle. Therefore the main task of the plugin program is to convert the point occupancy to the percentage occupancy.

Table 6-3 shows the result obtained with these occupancy states in Q-Learning Agents (both look up table and CMAC) and also with various fixed diversion percentages. It can be seen that Q-Learning agent has performed better than all the fixed diversion scenarios with CMAC giving better performance than the look-up-table approach. Therefore, in field implementation, the Q-Learning agent with CMAC function approximation seems to be most suitable as it has the generalization capabilities.

Table 6-3 Result of Simulation with Q-Learning Agents and Various Fixed Diversion Percentages (Using Occupancy States)

	Percentage Diversion							
	0%	20%	40%	60%	80% *(Base Case)*	100%	Q-Learning Look UP Table	Q-Learning CMAC
Mean Travel Time (sec)	720.4	635.7	547.5	516.1	568.8	746.5	491 *(13.7%)*	438.5 *(22.9%)*
Mean Speed (km/hr)	27.4	31.8	37.3	38.4	33.2	24.5	41.5 *(25%)*	47.1 *(41.9)*

Figures 6-3 and **6-4** respectively give the comparison of mean travel time and mean speed for Q-Learning agents and also with various static diversions. From the figures, it can be concluded that a Q-Learning agent incorporated with a function approximation technique such as CMAC is best suited to effectively control the traffic in a work zone area.

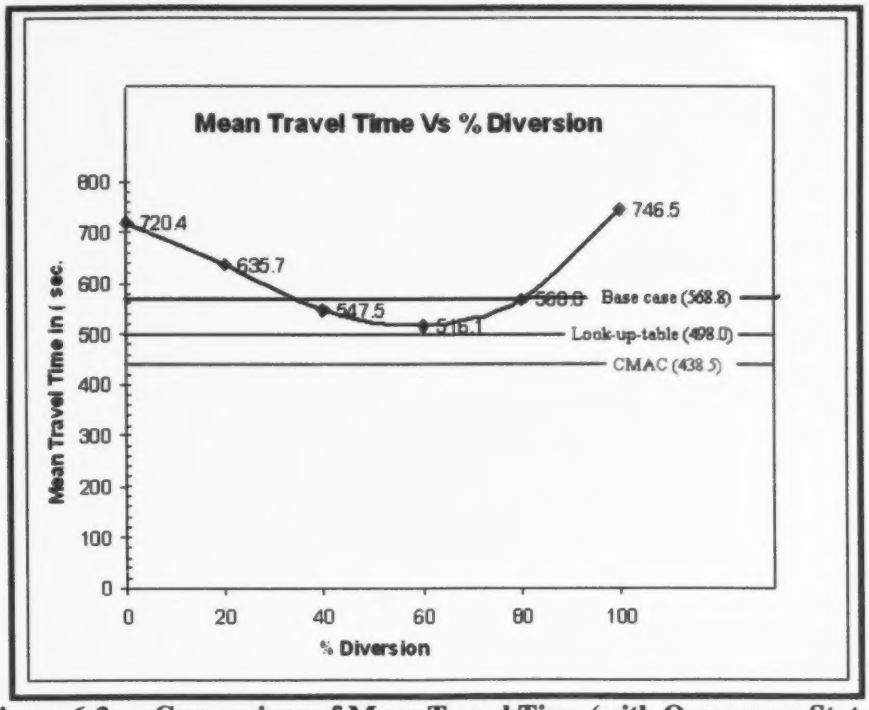


Figure 6-3 Comparison of Mean Travel Time (with Occupancy States)

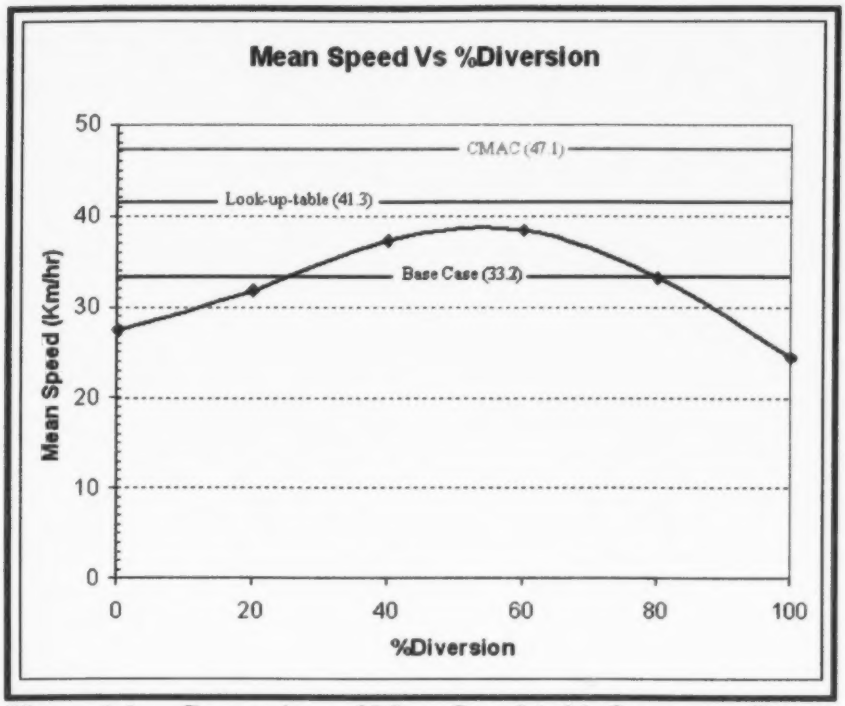


Figure 6-4 Comparison of Mean Speed (with Occupancy States)

7 FIELD TRIAL

After development of the software component of the TMS-Can system and obtaining promising results from simulated conditions, next step of this assignment was implementation of the system to real time traffic in the actual work zone at the trial site. The objective of the field trial was to obtain proof of the concept in the real world.

Field trials were scheduled to be undertaken in November 2005. Logistics for deployment of VMS signing and final verification of field data feeds from RTMS detectors were carried out at the trial site. Unexpectedly, during the week of November 7th, 2004 the barriers that had occupied the road at the trial site were removed and the area was opened to traffic. The termination of the construction activity and the corresponding lane reductions that were causing congestion on the road terminated the opportunity for the scheduled field trials in November 2004.

Since development for the TMS-Can system that had taken place in the laboratory was linked to the specific characteristics of the field site, attempts to delay the field trials until the return of construction activity in the spring of 2005 were assessed. It was determined, however, that sufficient activity would not be scheduled for the location in the future, and the only alternative would be to select an alternate site, including a redevelopment of the simulation, and schedule the field trial for a later date.

Several attempts were undertaken in 2005 to identify alternate trial sites and seek to schedule field trials. It was concluded in late 2005 that identification of an alternate would not be possible in 2005, and that the project would be concluded without completion of the fiend trials. The lessons learned from these events are discussed later in this report.

Although the developed system was not implemented in the field, this section describes the steps necessary to achieve the implementation of the system in the field at some point in the future.

In the next section, description of performance measures that could be used for evaluation of the TMS-Can is highlighted. At the end of this chapter, the lessons learned from this assignment and issues related to the system implementation is described.

7.1 Necessary Steps for Field Implementation of TMS-Can

The TMS-Can uses an artificial intelligence based computer agent to provide real time traffic control on work zones. The Q-Learning algorithm has proved to be performing well under simulated conditions. The reliability of the approach can only be confirmed with the field implementation of the agent in controlling the actual test site. The main requirement for implementing the developed system in the real world scenario is the

detector data. On receiving the data the aggregated values like average speed and average occupancy during 5 minutes and the outflow during each 1-minute period have to be formed for each duration. Based on these input values the agent would select the appropriate diversion percentage. This is to be displayed as appropriate message sign.

For this assignment, it was decided that the VMS should be located near Leslie intersection so as to provide enough time for the drivers to act on it. The messages should be so selected that the required diversion is possible. As per the Learning agent simulated, there should be room for displaying 6 different messages. The message should be selected to obtain the required diversion as much as possible. Following steps should be followed for the implementation of the control in the field.

- The VMS should be placed and set at the Leslie intersection to display the diversion message every 5-minute interval;

- The detectors on the Collector and Express should be finally checked for their reliability and accuracy. The detectors for collecting occupancy are given in **Table 7-1** on Collector and Express and those used for collecting outflow (this is required to evaluate the system performance) are shown in **Table 7-2**. The outflow data are used for the evaluation stage;

Table 7-1 Detectors Used for Collecting Occupancy Data

Detector ID	Location
a) Collector	
401DE0110DWC	W of Yonge CL
401DE0130DWC	E of Yonge CL
401DE0160DWC	E of Bayview CL
401DE0170DWC	E of Bayview CL
b) Express	
401DE0110DWE	W of Yonge EL
401DE0130DWE	E of Yonge EL
401DE0160DWE	E of Bayview EL
401DE0170DWE	W of Leslie EL

Table 7-2 Detectors Used for Collecting Outflow

Detector ID	Location
401DE0110DWC	W of Yonge CL
401DE0110DWE	W of Yonge EL
401DE0450DNR	W 401 - Bayview N RX
401DE0400DNR	E 401 - Yonge ST. RX

- The above data would be obtained from the ICAT (ITS Centre and Testbed at University of Toronto) database and platform (http://128.100.217.76:18800/icatdb/) that provides access to the live traffic data from MTO and City of Toronto;

- The online data (20 second duration) would be aggregated every 5-minutes in order to obtain the average occupancy on Collector and Express from the above mentioned detectors and will be stored in a file for ready access;

- The software agent would access the above average occupancy values and based on that it selects the appropriate diversion percentage based on the knowledge it has received during the simulated training;

- Each percentage diversion would be linked to a particular message on the VMS;

- The person managing the computer control would convey this message to the operator of the VMS display to activate the correct message. (During the test stage, it should be decided to convey the control action manually through the use of telephone. However long time implementation warrants automatic control transfer to apply the control in real time without any delay.); and

- The implementation of this project would be carried out for several morning peak hours.

7.2 Evaluation Criteria

Since the evaluation of the real-time advanced warning and traffic control system for its effectiveness is an important part of this project, several criteria were selected in order to evaluate performance of TMS-Can. The evaluation process will assess the problems. The

evaluation criteria will identify the measure of effectiveness for before and after scenarios of system usage.

The number of drivers diverted from Collector into the Express could be compared with the percentage expected to obtain the compliance rate of each message at the site. Similarly the time weighted outflows with and without implementation could be compared to obtain the goodness of the approach in reducing delay and improving travel time. The following data would be collected for evaluation:

- Number of drivers diverted out of that coming at the diversion point to obtain actual diversion happened;
- Outflow from each outflow detectors every minute to calculate the time-weighted outflow to act as a proxy for delay reduction. These values will be obtained for several morning peak hours with and without the implementation;
- Compare the actual diversion obtained and diversion expected by the message as a percentage deviation from what is expected; and
- Obtain the percentage change in delay.

7.3 Lessons Learned

Although site characteristics of work zones are expected to vary from site to site, there were a number of lessons learned that were common to most, if not all, of the system.

- *Construction site characteristics:* It is necessary that the construction site has primary and alternate routing available and that the site be one where congestion, linked to work zone activity, was actively taking place. The site recommended by MTO provided the right physical configuration (express/collector road network). It was found that identification of construction sites with the desired characteristics was difficult. Because of the desire to achieve a proof-of-concept of the TMS-Can system, it was particularly important in this project to have physical elements that allowed controlled measurement of results. This need significantly limited the available choice of sites, something that would not necessarily be at issue in real world deployment, after the concept has been proven.
- *Receiving uninterrupted and accurate data from the detectors*: This issue became a significant concern of this study. Work zones, by their nature, are temporary and involved major reconstruction of the road section. In locations where in-road traffic sensing equipment (loops) is used, it is not uncommon that most of the detectors or loops are not available, damaged or removed by construction. In the case of the original trial site and the potential alternatives that were explored, identification of sites that had field data capabilities, and would be able to

maintain that information input through the construction period, proved time consuming and difficult. . Therefore receiving accurate data from the work zone is highlighted as a challenging task that must be granted sufficient time and budget to ensure a successful trial.

- *Stakeholder agency involvement is critical:* This project worked with in-field locations as the trial site, both for the development of the models and the simulated testing carried out in the lab. The goal of proceeding to a field trial involves the ongoing assistance of stakeholder agencies. The goal of achieving a field trial of the concepts without impacting in-field operations proved difficult. It is critical that the project fully involve stakeholders, including all stakeholder participants who can have an impact on the outcome of the project, so as to achieve smooth transitions from lab to field. While stakeholders involved in this assignment were of significant assistance in allowing the assignment to progress as well as it did, the failure of achieving the field trial attests to the fact that sufficient coordination of stakeholder actions was not achieved by the project team. Coordination with all stakeholders is a primary issue that must be incorporated both in developing and implementing an ITS work zone system.

The failure to achieve deployment of the concepts developed and simulated in the laboratory setting highlighted a significant issue with the system, which it was highly dependent on the site selected for application and takes longer to develop for a selected site than would be acceptable for a commercial application. As noted, the system is developed and specifically designed around the physical characteristics of a given site. When the site for which a system had been developed was eliminated, quick deployment to an alternate was not simple to achieve.

The study was specifically designed to work with an existing construction work zone, using the parameters of that location and the information available from it, such as traffic count data, to build the simulation which then "learned" in the lab prior to field deployment. When that location became unavailable, the effort completed for the development at the site was also lost. Ideally, preparation of the TMS-Can system would be achieved in parallel with the planning and scheduling of a specific construction activity, as opposed to identifying a candidate from existing work underway.

It is also recognized that if the TMS-Can approach is to be successful it must be capable of being developed, simulated and field-deployed with significantly less effort that was measured in this study. While the authors are optimistic that this is possible, further study is needed in this area. To be successful as a technology that can be commercially deployed, TMS-Can must be able to be deployed rapidly and over a wider range of sites.

Finally, stakeholders involved in the project were of significant assistance in aiding the project team in understanding the realities of data availability and accuracy for the project. While more plentiful and more accurate data was desired by the project team, realities of traffic operations systems and construction activities resulted in less than ideal

conditions. As system development continues effort should be made to use available data as opposed to attempting to achieve unrealistic levels of information.

Also identified was the reality of externally-managed construction activity. This situation, where the road authority out-sources responsibility for construction management activity, presents a significant hurdle if coordinated activity between the road authority, the contract manager, and work zone management efforts are to be achieved. Effort to coordinate all parties for further development of the TMS-Can system would be needed to achieve the success of completing field trials.

To be truly successful as a commercial product TMS-Can would need to be capable of rapid deployment with essentially "plug and play" simplicity. The existing systems described in this report identify that the concept is viable and offers potential benefits, but significant work is required to simplify the system.

Along with efforts to complete field trials work should be undertaken to directly link vehicle detector deployment with the installation of a TMS-Can system not relying on others to provide the necessary data. Efforts should also be made to deploy the system on varying types of sites to ensure commercial viability.

SUMMARY AND CONCLUSIONS

The main goal of this research study was to develop functional and system requirements for an integrated, real-time temporary conditions traffic management system for work zones on Canadian roadways. The developed system, TMS-Can, uses a combination of electronic sensors, software, wireless and cabled communications networks, and electronic signs and variable message signs to manipulate traffic flow. The system is used in work zones to minimize highway disruptions and increase work zone safety.

The study area selected for the project is the construction site on the westbound collector of highway 401 west of Yonge Street at Hogg's Hollow. The construction reduced the collector to one lane at the site, which will greatly affect the collector capacity. There was a collector to express transfer upstream of the location, between Bayview Avenue and Leslie Street and an express to collector transfer west of the site, between Yonge and Avenue Road.

The research used a control strategy driven by Reinforcement Learning (RL) to provide traffic management capabilities and traveler information on traffic routing, and detours to drivers. Reinforcement Learning Technique from the Machine Learning community is utilized to develop the computer agent. Reinforcement Learning is an approach whereby the control agent directly learns to map sensed system states to optimal actions. A micro simulation tool - Paramics - was used to train the agent in an offline mode in a simulated environment.

The simulation tests have proved beyond doubt that a real time dynamic control like the Q-Learning approach can provide much better control than that of the fixed diversion. Even though both look-up table and CMAC approaches have performed well during simulation (with a better performance with CMAC), site implementation may warrant an approach with function approximation. This is because of the fact that during implementation, the agent may come across new states, which it has never visited during learning. Therefore the results obtained from the CMAC approach with occupancy state values are recommended for field implementation. During future site tests, the learned weights will be utilized to select action with highest Q-value. Online learning is also possible starting from the already learned weights to modify it as per the rewards obtained from the site. The learning rate and exploration rate have to be adjusted properly to obtain better results.

It is recommended that the developed system be implemented in an actual work zone at a trial site in order to achieve proof of concept in the real world. However, due to the temporary nature of work zones obtaining real-time data is a challenge. Potential improvements in this strategy could be made by starting the process earlier, and coordinating the use of TMS-Can as a part of a project, from the earliest stages. Such a planning effort would ensure availability of necessary site requirements and traffic data as well as ensure coordination of the trial with the construction project.

8 REFERENCES

Albus, J.S., A New Approach to Manipulator Control: The Cerebellar Model Articulation Controller (CMAC), Transactions of the ASME, 1975.

Canadian Strategic Highway Research Program, "Strategic Highway Research Program, Maintenance Work Zone Safety Devices-Development and Evaluation", Technical Brief, February 1995.

Darken, C., Chang, J., and Moody, J. Learning Rate Schedules for Faster Stochastic Gradient Search, in Neural Networks for Signal Processing II, Kung, Fallside, Sorenson, and Kamm, eds., IEEE, Piscataway, New Jersey, 1992.

EMME/2 User's Guide Manual Release 9, INRO Consultants INC., 1998

FHWA, "Intelligent Transportation Systems in Work Zones A Cross-cutting Study", November 2002.

FHWA, "Safety Applications of ITS in Rural Areas", http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13609.html, September 2002.

Gasser, M., Computation in Cognitive Science, Indiana University, www.Indiana.edu/~q320/Notes/rl-explore.html, Accessed in November 2004.

Glanz, F.H, Miller, W.T., and Kraft, L.G., An overview of the CMAC neural network, IEEE Conference on Neural Networks for Ocean Engineering, Washington, DC, pp. 301-308, 1991.

ICAT (ITS Centre and Testbed) data base and platform, University of Toronto http://128.100.217.76:18800/icatdb/.

Jacob, C., and Abdulhai, B., Seamless Multi-Jurisdictional Traffic Corridor Control Using Reinforcement Learning: Theoretical Framework and Early Results. 83rd Annual Conference, Transportation Research Board, 2004.

Kaelbling, L. P., Littman, M.L., Moore, A. W., Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, Vol.4, pp.237-285, 1996.

Klashinsky, R. and Bushman, R., "Dynamic Work Zone Safety System", Annual Conference of the Transportation Association of Canada, Edmonton, Alberta, 2000.

Maze, T, Kamyab, A., Poole, C., Meyer, E., McCoy, P., Pesti, G., Shaik, N, Bernhardt, L.S., and Virkler, M. R.," Midwest States Smart Work Zone Deployment Initiative: Technology Evaluations – Year1", http://www.matc.unl.edu/project/, May 2000.

McCoy, P., Pesti, G., and Byrd, P.S., "Alternative Driver Information to Alleviate Work Zone Related Delays", Final Report, NDOR Research Project Number SPR-PL-1 (35) P513, Nebraska Department of Roads, February 1999.

Millar, G., The Application of Paramics Microscopic Simulation in Solving Real Transportation Problems, Paramics Traffic Simulation, Third Asia-Pacific Conference 1999.

Miller, W. T., and Ghanz F. H., The University of New Hampshire Implementation of the Cerebellar Arithmetic Controller-CMAC, Robotic Laboratory, University of New Hampshire, Durham, New Hampshire, 1996.

Mitchell, T., Machine Learning. McGraw-Hill. New York, 1997.

Pant, P.D., Polycarpou, M.M., Sankaranarayanan, P., Li, B., Hu, X., and Hossain, A.B., "Travel Time Prediction System (TIPS) for Freeway Work Zones", Traffic And Transportation Studies Proceedings of ICTTS '98, 1998

Pant, P.D., "Smart Work Zone Systems", Pant, http://ops.fhwa.dot.gov/wz/workshops/accessible/Pant_paper.htm, WORKZONE MOBILITY AND SAFETY PROGRAM, FHWA.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y., Review of Road Traffic Control Strategies. Proceedings of the IEEE, Vol.91, No.12, December 2003.

Quadstone Ltd., Modeller and Programmer V3.0 User Guide and Manual. 2000, Quadstone Limited., Edinburgh, UK. http://www.paramics-online.com

SRF Consulting Group, Inc., "Portable Traffic Management System Smart Work Zone Application Operational Test Evaluation Report", Minnesota Department of Transportation, May 1997.

Smart, W.D., Making Reinforcement Learning Work on Real Robots, PhD Thesis. Department of Computer Science, Brown University, Providence, Rhode Island 2002.

Sutton, R.S., Generalization in Reinforcement Learning: Successful Examples Using Sparse Coarse Coding, Advances in Neural Information Processing Systems 8, pp. 1038-1044, MIT Press, 1996

Sutton, R.S., and Barto, A.G., Reinforcement Learning-An Introduction. MIT Press, Cambridge, Massachusetts, 1998.

Thommana, J.V. and Sivanandan, R., "Real-time Advanced Warning And Traffic Control Systems for Work Zones: Examination of Requirements And Issues", Master of Science in Civil Engineering, Virginia Polytechnic Institute and State University, 1997.

Tudor, L.H., Meadors, A. and Plant, R., " Deployment of Smart Work Zone Technology in Arkansas", Transportation Research Board 2003 Annual Meeting CD-ROM, Washington D.C., 2003.

Watkins, C.J.C.H. Learning from delayed Rewards. Ph.D. Thesis. Kings College, University of Cambridge, Cambridge, UK, 1989.

APPENDIX
Photographic Log – Site Visit

Figure A-1 Collector with 4 Lanes East of Leslie Street on-Ramp Interchange

Figure A-2 East of Transfer Point from Collector to Express, West of Leslie Street

Figure A-3 Collector Reduces to 3 Lanes After The Transfer, East of Bayview Avenue

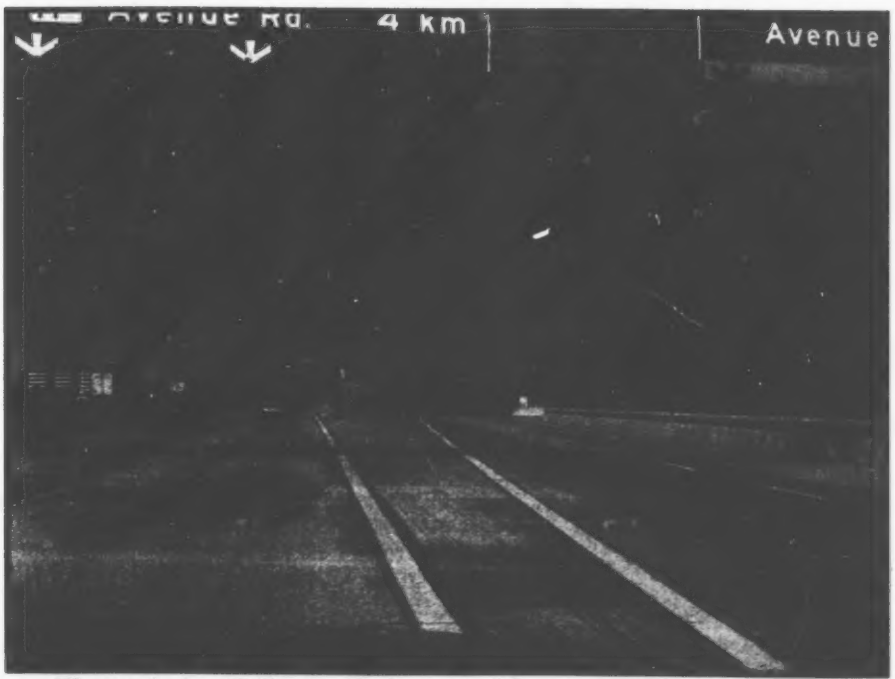


Figure A-4 Collector Reduces to 2 Lanes, at Bayview Avenue

Figure A-5 Collector Reduces to a Single Lane, at Yonge Street